WSGR Docket No. [26050-709.501]

PATENT APPLICATION

HUMANIZED ANTIBODIES AGAINST VASCULAR ENDOTHELIAL GROWTH FACTOR

Inventor(s): Pingyu ZHONG
Citizen of the United States, Residing at
1475 Tyler Parkway
Mountain View, CA 94040

Peizhi LUO
Citizen of the United States, Residing at
1671 Canary Drive
Sunnyvale, CA 94087

Kevin C. WANG
Citizen of the United States, Residing at
2316 20th Avenue
San Francisco, CA 94116

Mark HSIEH
Citizen of the United States, Residing at
1557-D Pershing Drive
San Francisco, CA 94129

Yan LI
Citizen of China, Residing at
3231 Woodmont Drive
San Jose, CA 95118

Assignee: Abmaxis, Inc.
453 Ravendale Drive
Mountain View, CA 94043

WSGR

Wilson Sonsini Goodrich & Rosati
PROFESSIONAL CORPORATION

650 Page Mill Road
Palo Alto, CA 94304
(650) 493-9300

Express Mail Label No. EV 333 491 597 US

HUMANIZED ANTIBODIES AGAINST VASCULAR ENDOTHELIAL GROWTH FACTOR

CROSS REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of U.S. Application No. 10/443,134, filed May 20, 2003, entitled "Generation and Selection of Protein Library in Silico" which is a continuation-in-part of U.S. Application No. 10/153;159, filed May 20, 2002, entitled "Structure-Based Selection And Affinity Maturation of Antibody Library, and also a continuation-in-part of Application No. 10/153,176, filed May 20, 2002, entitled "Generation Affinity Maturation of Antibody Library in Silico", both of which are a continuation-in-part of U.S. Patent Application Serial No: 10/125,687 entitled "Structure-based construction of human antibody library" filed April 17, 2002, which claims the benefit of U.S. Provisional Application Serial No: 60/284,407 entitled "Structure-based construction of human antibody library" filed April 17, 2001. These applications are incorporated herein by reference.

BACKGROUND OF THE INVENTION

The present invention provides antibodies against human antigens and methods for generating high affinity antibodies against these targets. More particularly, the invention provides humanized or human antibodies against vascular endothelial growth factor (VEGF) and methods for generating such anti-VEGF antibodies. In addition, the invention provides compositions, kits and methods of using these antibodies and derivatives thereof to inhibit angiogenesis in vitro, and for diagnosing or treating diseases associated with abnormal angiogenesis such as cancer, rheumatoid arthritis, ischemic-reperfusion related brain edema and injury, cortical ischemia, ovarian hyperplasia and hypervascularity, endometriosis, psoriasis, diabetic retinopaphy, and other ocular angiogenic diseases.

Angiogenesis has been involved in many physiological and pathological processes. Angiogenesis consists of multiple steps that ultimately resulting in

proliferation and differentiation of endothelial cells, and formation of tubes and cavities (angiogenesis). Factors that promote angiogenesis include VEGF, aFGF, bFGF, TGF-α, TGF-β, HGF, TNF-α, angiogenin, IL-8, etc., whereas factors that inhibit angiogenesis include thrombospondin (Good et al. *Proc. Natl. Acad. Sci. USA.* 87:6624-6628 (1990)), the N-terminal fragment of prolactin (Clapp et al. *Endocrinology,* 133:1292-1299 (1993)), kringle 5 domain of plasminogen (Cao et al., J. Biol. Chem., 271:29461-29467 (1996), angiostatin (O'Reilly et al. *Cell,* 79:315-328 (1994)) and endostatin (O'Reilly et al. *Cell,* 88:277-285 (1996)).

Vascular endothelial cell growth factor (VEGF) is a growth factor acting specifically with its receptors on vascular endothelial cells to promote their angiogenesis. It is known that angiogenesis plays an important role in the development of new vasculature from preexisting blood vessels and/or circulating endothelial stem cells (Asahara et al., Science, 275(5302):964-967, 1997; Springer et al., Mol. Cell, 2(5):549-558, 1998; Folkman and Shing, J. Biol. Chem., 267:10931-10934, 1992). Angiogenesis also plays a vital role in many physiological processes, such as embryogenesis, wound healing and menstruation. More importantly, angiogenesis is further involved in pathological conditions such as tumor formation, metastasis, diabetic retinopathy, etc. It is known that the growth of a solid tumor requires tumor vascularization for supplying oxygen and nutrients and the metastasis of tumor cells occurs through blood vessels resulting from the tumor vascularization. VEGF is believed to be a pivotal angiogenic factor in this vascularization for tumors. Therefore, it is expected that the growth and metastasis of tumor can be inhibited by certain substances neutralizing the vascularization activity of VEGF. Recent studies (Burrows and Thorpe, Pharmacol. Ther., 64:155-174, 1994; Proc. Natl. Acad. Sci. USA, 90:8996-9000, 1994) have used such a strategy to target the vasculature of solid tumors. Targeting the blood vessels of the tumors, rather than the tumor cells themselves, has certain advantages in that it is not likely to lead to the development of resistant tumor cells, and that the targeted cells are readily accessible. Moreover, destruction of the blood vessels leads to an amplification of the anti-tumor effect, as many tumor cells rely on a single vessel for their oxygen and nutrients

In nearly half of diabetics diabetic retinopathy occurs as one of complications of diabetes. It is believed that the formation of microcapillaries is promoted in diabetic retinopathy by oxygen deficiency. These microcapillaries will sooner or later be ruptured to bleed to form scar tissue, leading to detached retinas. Age-related macular degeneration is another eye disease that has been demonstrated to be involved in pathological vascularization in the retina. Hence, it is expected that inhibition of vascularization can prevent retinopathy from developing. Based on experiments using monkey, Miller et al. reported that VEGF is related very closely to the development of vegetative retinopathy (Miller et al.: Am. J. Pathol. 145, 574-584 (1994)). For this reason, a substance neutralizing the vascularization activity of VEGF is considered useful for preventing or treating diabetic retinopathy and AMD (Lopez et al. Invest. Ophtalmo. Vis. Sci. 37:855-868 (1996)).

The recognition of VEGF as a primary stimulus of angiogenesis in pathological conditions has led to various attempts to block VEGF activity. Inhibitory anti-VEGF receptor antibodies, soluble receptor constructs, antisense strategies, RNA aptamers against VEGF and low molecular weight VEGF receptor tyrosine kinase (RTK) inhibitors have all been proposed for use in interfering with VEGF signaling (Siemeister et al., Cancer Metastasis Rev., 17(2):241-248., 1998). In fact, monoclonal antibodies against VEGF have been shown to inhibit human tumor xenograft growth and ascites formation in mice (Kim et al., Growth Factors 7:53 (1992); Nature 362:841-844 (1993); Asano et al., Hybridoma, 17:185-90, (1998); Mesiano et al., Am. J. Pathol., 153(4):1249-1256, (1998); Luo et al., Cancer Res., 58(12):2594-2600, (1998); Cancer Res., 58(12):2652-2660, (1998); Borgstrom et al. Cancer Res. 56:4032-4039 (1996); Borgstrom et al., Prostate, 35(1):1-10, (1998)). Furthermore, the same strategy using anti-angiogenic molecules, anti-VEGF antibody, and VEGF antagonists have been utilized in experimental treatment of AMD and diabetic retinopathy (Adamis et al. Arch. Ophthalmol. 114:66-71 (1996)). For the therapeutic applications, antibodies are generally engineered to reduce their toxicities in repeated dosage by humanization, if they are derived originally from mouse and to improve other attributes such as binding affinity with the target molecules by affinity maturation (Winter and Milstein, Nature, 349:293-299, (1991); Baca et al, J. Biol. Chem., 272(16):10678-84, (1997); Presta, et al.,

Cancer Res., 57:4593-4599, (1997); Chen et al. (1999) J. Mol. Biol. 293:865-881; and Ryan et al. (1999) Toxicologic Pathology, 27(1):78-86).

Although the foregoing studies underscore the importance of VEGF in solid tumor growth, and its potential as a target for tumor therapy, the identification of additional agents that inhibit VEGF-induced angiogenesis would be of benefit in expanding the number of therapeutic options. The development of therapeutic agents that specifically inhibit VEGF to bind with its receptor represents important alternatives to target angiogenesis more effective with potentially improved therapeutic benefits.

SUMMARY OF THE INVENTION

The present invention provides an innovative methodology for engineering proteins with desired structures and functions, especially for engineering antibodies with desirable properties from a therapeutic perspective, including high binding affinity for the target antigen, ability to effectively inhibit abnormal cell proliferation in vitro and in vivo, and minimal toxicity or side effects.

In one aspect of the invention, methods are provided for designing and selecting antibodies against human antigens with high affinity and specificity in silico and in vitro. In some particular embodiments, methods are provided for designing and selecting humanized or fully human antibodies against vascular endothelial growth factor (VEGF) with high affinity and specificity.

In another aspect of the invention, monoclonal antibodies against VEGF are provided. In particular, humanized or human anti-VEGF monoclonal antibodies are provided that can bind to human VEGF with high affinity. Preferably, these antibodies can inhibit VEGF-induced proliferation of endothelial cells in vitro and inhibit VEGF-induced angiogenesis in vivo. These antibodies and their derivative can be used in a wide variety of applications such as diagnosis, prevention, and treatment of diseases such as cancer, AMD, diabetic retinopathy, and other diseases derived from pathological angiogenesis.

In one embodiment, a monoclonal antibody is provided that specifically binds to a human VEGF with dissociation constant K_d equal to or lower than 0.2 nM, optionally lower than 0.1 nM, optionally lower than 0.08 nM, optionally lower than 0.05 nM, optionally lower than 0.01 nM, or optionally lower than 0.005 nM, in the form of scFv, Fab, or other form of antibody measured at a temperature of about 4°C, 25°C, 37°C or 42°C.

Preferably, K_d of the anti-VEGF antibody, if in the form of scFv and measured at a temperature of 35-37°C, is lower than 1 nM, optionally lower than 0.8 nM, optionally lower than 0.5 nM, optionally lower than 0.2 nM, optionally lower than 0.1 nM, optionally lower than 0.08 nM, optionally lower than 0.05 nM, optionally lower than 0.01 nM, or optionally lower than 0.005 nM.

Also preferably, K_d of the anti-VEGF antibody, if in the form of Fab and measured at a temperature of 35-37°C, is lower than 0.5 nM, optionally lower than 0.2 nM, optionally lower than 0.1 nM, optionally lower than 0.08 nM, optionally lower than 0.05 nM, optionally lower than 0.01 nM, or optionally lower than 0.005 nM.

In inhibition of VEGF-induced proliferation of endothelial cells in vitro, the the anti-VEGF antibody preferably has an effective dose for inhibition of 50% cell proliferation ED_{50} equal to or lower than 10 nM, optionally lower than 5 nM, optionally lower than 1 nM, optionally lower than 0.5 nM, optionally lower than 0.1 nM, optionally lower than 0.05 nM, or optionally lower than 0.01 nM in the form of scFv, Fab or other form of antibody.

In a preferred embodiment, a monoclonal antibody is provided that specifically binds to a human VEGF and has V_L comprising the amino acid sequence of $X_1X_2X_3X_4TQX_5PSX_6X_7SX_8X_9X_{10}GX_{11}X_{12}X_{13}X_{14}IX_{15}C\underline{X_{16}X_{17}SX_{18}X_{19}IX_{20}X_{21}X_{22}X_{23}X_{24}}$ $WYQQX_{25}PGX_{26}APX_{27}X_{28}LX_{29}Y\underline{X_{30}X_{31}X_{32}X_{33}LX_{34}X_{35}}GVX_{36}X_{37}RFSGX_{38}X_{39}SGTDF$ $X_{40}LTIX_{41}X_{42}LQX_{43}X_{44}DX_{45}AX_{46}YYC\underline{QQX_{47}X_{48}X_{49}X_{50}PX_{51}T}FGX_{52}GTKX_{53}X_{54}IK$, wherein the underlined regions are designated as V_L/CDR1, V_L/CDR2, and V_L/CDR3, respectively, whereas the rest of the region is designated as framework, and wherein X_1 is D, E or A; X_2 is I, or T; X_3 is V, E, K, R, Q, or T; X_4 is M, or L; X_5 is S, or T, X_6 is S, or T; X_7 is L, or V; X_8 is A, or V; X_9 is S, or T; X_{10} is P, V, L, A, or I; X_{11} is E, or D; X_{12} is R, or T; X_{13} is A, or V I; X_{14} is T, or A; X_{15} is T, S, or A; X_{16} is S, R, N, K, H, or Q; X_{17} is A, or S; X_{18} is Q, or R; X_{19} is S, D, A, or P; X_{20} is S, G, R, T, or Y; X_{21} is T, N, S, D, or K; X_{22} is Y, or D; X_{23} is L, or I; X_{24} is A, N, or T; X_{25} is K, or I; X_{26} is Q, K, T, or I; X_{27} is R, K, Q, N, H, S, or E; X_{28} is V, or L; X_{29} is I, or V; X_{30} is F, A, G , D, or S; X_{31} is A , or T; X_{32} is S, or T; X_{33} is N, S, R, or T; X_{34} is A, H, or Q; X_{35} is S , or G; X_{36} is P, T; X_{37} is S, N, D, G, or Y; X_{38} is S, or T; X_{39} is G, or R; X_{40} is T, or A; X_{41} is S, or R; X_{42} is S, or R; X_{43} is P, or A; X_{44} is E , or D; X_{45} is F, V, or S; X_{46} is V, T, I , A , or S; X_{47} is Y, or S; X_{48} is S, Y, or N; X_{49} is S , or T; X_{50} is T, V, A, P, K, G, S, or I; X_{51} is W , or Y; X_{52} is Q , or G; X_{53} is V, or L; and X_{54} is E, D , or A.

Such preferred V_L sequences may be combined with the preferred V_H sequences or V_H of other antibodies, provided that the antibody so produced binds to the human VEGF with a desired affinity.

In another preferred embodiment, a monoclonal antibody is provided that specifically binds to a human VEGF and has V_L comprising the amino acid sequence of $X_1X_2X_3$LTQPPSX_4SX_5TPGQX_6VTISCS$\underline{GX_7X_8SNX_9GX_{10}NX_{11}VX_{12}}$WYQQ$X_{13}PGX_{14}$APKK$X_{15}LX_{16}Y\underline{X_{17}NX_{18}X_{19}RPS}GVPX_{20}RX_{21}SGSX_{22}SX_{23}$TSASLAISGL$X_{24}$SEDEADYYC$\underline{X_{25}X_{26}WDDSLX_{27}GYV}FGX_{28}GTX_{29}$LTVL, wherein the underlined regions are designated as V_L/CDR1, V_L/CDR2, and V_L/CDR3, respectively, whereas the rest of the region is designated as framework, and wherein X_1 is Q L, or N; X_2 is P A F, or S; X_3 is V, or M; X_4 is A, or T; X_5 is G, or A; X_6 is R, or S; X_7 is S, or T; X_8 is S, T Y, or N; X_9 is I , or V; X_{10} is S, or R; X_{11} is S, P, N, A, or T; X_{12} is N, T , or Y; X_{13} is L, or F; X_{14} is T, or A; X_{15} is V, L, or F; X_{16} is M, or I; X_{17} is G, T, or S; X_{18} is N, or D; X_{19} is Q, or E; X_{20} is D, or E; X_{21} is F, or L; X_{22} is K, or R; X_{23} is G, or A; X_{24} is Q, L, or R; X_{25} is A, or G; X_{26} is A, S, or T; X_{27} is N, S, or T; X_{28} is T, or A; and X_{29} is K, or Q.

Such preferred V_L sequences may be combined with the preferred V_H sequences or V_H of other antibodies, provided that the antibody so produced binds to the human VEGF with a desired affinity.

In yet another preferred embodiment, a monoclonal antibody is provided that specifically binds to a human VEGF and has V_L comprising the amino acid sequence of QSALTQPPSVSGAPGQRVTISCTGRSSNIGAGHDVHWYQQLPGTAPKLLIYANDQRPSGVPDRFSDSKSGTSASLGISGLRSEDEADYFCATWDDSLHGYVFGTGTKVTVL (SEQ ID No: 54). This V_L sequence may be combined with the preferred V_H sequences or V_H of other antibodies, provided that the antibody so produced binds to the human VEGF with a desired affinity.

In yet another preferred embodiment, a monoclonal antibody is provided that specifically binds to a human VEGF and has V_H comprising the amino acid sequence of X_1X_2QLVX_3SGGGX_4VQPGGX_5LRLX_6CAX_7S$\underline{GX_8X_9X_{10}X_{11}X_{12}X_{13}GX_{14}N}WX_{15}$RQAPGKG$X_{16}$EWVG$\underline{WX_{17}NTX_{18}X_{19}GX_{20}X_{21}TYX_{22}X_{23}X_{24}FX_{25}R}RX_{26}TX_{27}SX_{28}X_{29}X_{30}SKX_{31}X_{32}X_{33}YLQX_{34}$NSLRAEDTAVYYCA$\underline{X_{35}YPX_{36}YYGX_{37}SHWYFDV}WX_{38}$QGTLVTVSS, wherein the underlined regions are designated as CDR1, CDR2, and CDR3, respectively, whereas the rest of the region is designated as framework according to Kabat nomenclature, and wherein X_1 is E, or Q; X_2 is V, or G; X_3 is Q, or E; X_4 is V, or L; X_5 is S, or T; X_6 is S T, or R; X_7 is A, or V; X_8 is Y, or F; X_9 is T, D, N, S, or A; X_{10} is

F, or L; X_{11} is T, D, Y, A, S, or N; X_{12} is N, H, or S; X_{13} is Y, or F; X_{14} is M, L, I, or V; X_{15} is I, V, or L; X_{16} is L, or P; X_{17} is I, or V; X_{18} is Y, or N; X_{19} is T, or N; X_{20} is E, or A; X_{21} is P, T, or S; X_{22} is A, or V; X_{23} is A, H, Q, P, D, or E; X_{24} is D, or E; X_{25} is K, or T; X_{26} is V, F, or L; X_{27} is F, or I; X_{28} is L, or R; X_{29} is D, or N; X_{30} is T, or N; X_{31} is S, or N; X_{32} is T, Q, P, or K; X_{33} is A, V, or P; X_{34} is L, or M; X_{35} is K, or R; X_{36} is H, or Y; X_{37} is S, R, or T; and X_{38} is G, or A.

Such preferred V_H sequences may be combined with the preferred V_L sequences or V_L of other antibodies, provided that the antibody so produced binds to the human VEGF with a desired affinity.

In one embodiment, a monoclonal antibody is provided that specifically binds to a human VEGF and has V_L comprising the amino acid sequence selected from the group consisting of SEQ ID NOs:2-54, more preferably comprising the amino acid sequence selected from the group consisting of SEQ ID NO:14, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:44, SEQ ID NO:47, and SEQ ID NO:54.

Such preferred V_L sequences may be combined with the preferred V_H sequences or V_H of other antibodies, provided that the antibody so produced binds to the human VEGF with a desired affinity.

In one embodiment, a monoclonal antibody is provided that specifically binds to a human VEGF and has V_H comprising the amino acid sequence selected from the group consisting of SEQ ID NOs:57-110 and SEQ ID NOs:285-310, and preferably comprising the amino acid sequence selected from the group consisting of SEQ ID NOs:61-64, SEQ ID NO:67, 68, 70, 75, 83, 88, 89, 90, 91, 92, 93, 94, and 96-110.

Such preferred V_H sequences may be combined with the preferred V_L sequences or V_L of other antibodies, provided that the antibody so produced binds to the human VEGF with a desired affinity.

In yet another embodiment, a monoclonal antibody is provided that specifically binds to a human VEGF and has CDR1 in the V_L region (V_L/CDR1) comprising the amino acid sequence selected from the group consisting of SEQ ID NOs:164-194.

Such preferred V_L/CDR1 may be combined with other regions of preferred light chain, or of other light chain, and the preferred heavy chain variable region sequence or

with other heavy chain variable region sequences, provided that the antibody so produced binds to human VEGF with desired affinity.

In yet another embodiment, a monoclonal antibody is provided that specifically binds to a human VEGF and has CDR2 in the V_L region (V_L/CDR2) comprising the amino acid sequence selected from the group consisting of SEQ ID NOs:195-209.

Such preferred V_L/CDR2 may be combined with other regions of preferred light chain, or of other light chain, and the preferred heavy chain variable region sequence or with other heavy chain variable region sequence, provided that the antibody so produced binds human VEGF.

In yet another embodiment, a monoclonal antibody is provided that specifically binds to a human VEGF and has CDR3 in the V_L region (V_L/CDR3) comprising the amino acid sequence selected from the group consisting of SEQ ID NOs:210-228.

Such preferred V_L/CDR3 may be combined with other regions of preferred light chain, or of other light chain, and the preferred heavy chain variable region sequence or with other heavy chain variable region sequence, provided that the antibody so produced binds to human VEGF with desired affinity.

In yet another embodiment, a monoclonal antibody is provided that specifically binds to a human VEGF and has a framework region (FR) CDR3 in the V_L region (V_L/FR) comprising the amino acid sequence selected from the group consisting of: SEQ ID NO:229-269, and preferably comprising the amino acid sequence selected from the group consisting of SEQ ID NO:232, 235, 237, 251, 255, 263, and 265.

Such preferred V_L/FR may be combined with CDR regions of preferred light chain, or of other light chain, and the preferred heavy chain variable region sequence or with other heavy chain variable region sequence, provided that the antibody so produced binds to human VEGF with desired affinity.

In one embodiment, a monoclonal antibody is provided that specifically binds to a human VEGF and has CDR1 in the V_H region (V_H/CDR1) comprising the amino acid sequence of $GX_1X_2X_3X_4X_5X_6GX_7N$, wherein X_1 is Y, or F; X_2 is D, N, T, S, or A; X_3 is F, or L; X_4 is T, D, S, Y, A, or N; X_5 is H, N, or S; X_6 is Y, or F; X_7 is M, L, I, or V.

Further preferably V_H/CDR1 of the monoclonal antibody comprises the amino acid sequence selected from the group consisting of: SEQ ID NOs:111-135.

Such preferred V_H/CDR1 may be combined with other regions of preferred heavy chain or of other heavy chain, and the preferred light chain variable region sequence or with other light chain variable region sequence, provided that the antibody so produced binds human VEGF.

In one embodiment, a monoclonal antibody is provided that specifically binds to a human VEGF and has CDR2 in the V_H region (V_H/CDR2) comprising the amino acid sequence of $WX_1NTX_2X_3GEX_4TYX_5X_6X_7FX_8R$, wherein X_1 is I, or V; X_2 is Y, or N; X_3 is T, or N; X_4 is P, T, or S; X_5 is A, or V; X_6 is A, Q, P, H, D, or E; X_7 is D, or E; and X_8 is K, or T

Preferably V_H/CDR2 of the monoclonal antibody comprises the amino acid sequence selected from the group consisting of: SEQ ID NOs:136-156.

Such preferred V_H/CDR2 may be combined with other regions of preferred heavy chain or of other heavy chain, and the preferred light chain variable region sequence or with other light chain variable region sequence, provided that the antibody so produced binds to human VEGF with desired affinity.

In one embodiment, a monoclonal antibody is provided that specifically binds to a human VEGF and has CDR3 in the V_H region (V_H/CDR3) comprising the amino acid sequence of $KYPX_1YYGX_2SHWYFDV$, wherein X_1 is Y, or H, and X_2 is R.

Preferably, the anti-VEGF antibody V_H/CDR3 has the amino acid sequence selected from the group consisting of SEQ ID NOs:311-337.

Such preferred V_H/CDR3 may be combined with other regions of preferred heavy chain or of other heavy chain, and the preferred light chain variable region sequences or with other light chain variable domain sequences, provided that the antibody so produced binds to human VEGF with desired affinity.

In one embodiment, a monoclonal antibody is provided that specifically binds to a human VEGF and has FR in the V_H region (V_H/FR) comprising the amino acid sequence of

$X_1VQLVX_2SGGGX_3VQPGGX_4LRLX_5CAX_6S$/CDR1/$WX_7RQAPGKGLEWVG$/CDR2/ $RX_8TX_9SX_{10}DX_{11}SKX_{12}X_{13}X_{14}YLQX_{15}NSLRAEDTAVYYCA$/CDR3/$WX_{16}QGTLVTV$ SS, wherein X_1 is E, or Q; X_2 is Q, or E; X_3 is V, or L; X_4 is S, or T; X_5 is S, T, or R; X_6

is A, or V; X_7 is I, or V; X_8 is F, or V; X_9 is F, or I; X_{10} is L, or R is X_{11} is T, or N; X_{12} is S, or N; X_{13} is T, Q, or K; X_{14} is A, or V; X_{15} is M, or L; and X_{16} is G, or A.

In one embodiment, a monoclonal antibody is provided that specifically binds to a human VEGF and has a V_L and V_H pair selected from the group consisting of: SEQ ID NO:1 and 70; SEQ ID NO:1 and 67; SEQ ID NO:1 and 75; SEQ ID NO:1 and 83; SEQ ID NO:14 and 55; SEQ ID NO:1 and 101; SEQ ID NO:1 and 100; SEQ ID NO:14 and 102; SEQ ID NO:1 and 103; SEQ ID NO:1 and:104; SEQ ID NO:1 and 105; SEQ ID NO:36 and 100; SEQ ID NO:26 and 100; SEQ ID NO:28 and 100; SEQ ID NO:37 and 100; SEQ ID NO:44 and 100; SEQ ID NO:54 and 100; and SEQ ID NO:47 and 100, preferably selected from the group consisting of SEQ ID NO:28 and 61; SEQ ID NO:28 and 62; SEQ ID NO:28 and 63; SEQ ID NO:28 and 64; SEQ ID NO:28 and 68; SEQ ID NO:28 and 85; SEQ ID NO:28 and 86; SEQ ID NO:28 and 87; SEQ ID NO:28 and 88; SEQ ID NO:28 and 89; SEQ ID NO:28 and 90; SEQ ID NO:28 and 91; SEQ ID NO:28 and 92; SEQ ID NO:28 and 93; SEQ ID NO:28 and 94; SEQ ID NO:28 and 95; SEQ ID NO:28 and 96; SEQ ID NO:28 and 97; SEQ ID NO:28 and 98; SEQ ID NO:28 and 99; SEQ ID NO:28 and 106; SEQ ID NO:28 and 107; SEQ ID NO:28 and 108; SEQ ID NO:28 and 109; and SEQ ID NO:28 and 110.

The binding affinity of the above-described antibody to a human VEGF, represented by dissociation constant K_d, is optionally lower than 100 nM, optionally lower than 10 nM, optionally lower than 8 nM, optionally lower than 5 nM, optionally lower than 1 nM, optionally lower than 0.8 nM, optionally lower than 0.5 nM, optionally lower than 0.2 nM, optionally lower than 0.1 nM, optionally lower than 0.08 nM, optionally lower than 0.05 nM, optionally lower than 0.01 nM, or optionally lower than 0.005 nM, in the form of scFv, Fab, or other form of antibody measured at a temperature of about 4°C, 25°C, 37°C or 42°C.

The antibodies resulted from combination of the full-length V_H and V_L, V_H/CDR, V_H/FR, V_L/CDR, V_L/FR (e.g., the hit variants provided in Figures 1A and 1B and amino acid sequences shown in Figures 1C and 1D) disclosed in the present invention are within the scope of the present invention; and such combination does not include the anti-VEGF antibodies disclosed in US Patent application serial No: 09/056,160, publication No: 2002/0032315.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1A shows an AA-PVP profile of V_L hit variants designed by using the inventive methodology *in silico.*

Figure 1B shows an AA-PVP profile of V_H hit variants designed by using the inventive methodology *in silico.*

Figure 1C shows amino acid sequences of full-length V_L, V_L/CDR, and V_L/FR of certain embodiments of the antibodies according to the present invention.

Figure 1D shows amino acid sequences of full-length V_H, V_H/CDR, and V_H/FR of certain embodiments of the antibodies according to the present invention.

Figures 2A and 2B show the affinity analysis of a panel of variants of humanized anti-VEGF antibody using BioCore biosensor at 25°C. Antibody fragments in the format of single chain were expressed as described below and purified. The measurement was done by measuring the change of SPR units (y-axis) vs time (x-axis) when a purified antibody binds its antigen (VEGF) immobilized on the CM5 biochip at 25°C. Kd values were calculated by the ratio of K_{on} to K_{off}. A ratio of a variant Kd to that of hAB1 for each variant antibody was also calculated. K_{on} is defined as association rate constant expressed as the reciprocal of the multiple of concentration of the antibody fragment (in molar) and time (in second); whereas K_{off} is defined as dissociation rate constant expressed as the reciprocal of time (in second); Kd is defined as dissociation constant, which is the ratio of K_{off} to K_{on} expressed in molar concentration. hAB1 contains SEQ1 and SEQ55 (which, if in the form of Fab, would be essentially Fab-12 as described in Chen et al. (1999) J. Mol. Biol. 293:865-881); hAB2 contains SEQ1 and SEQ70; hAB3 contains SEQ1 and SEQ67; hAB4 contains SEQ1 and SEQ83; hAB5 SEQ1 and SEQ75; hAB9 contains SEQ14 and SEQ102; hAB10 contains SEQ1 and SEQ103; hAB11

contains SEQ1 and SEQ104; hAB12 contains SEQ1 and SEQ105, as V_L and V_H, respectively.

Figure 2C shows the affinity analysis of a panel of variants of humanized anti-VEGF antibody using BioCore biosensor at 35°C. Antibody fragments in the format of single chain were expressed as described below and purified. The measurement was done by measuring the change of SPR units (y-axis) vs time (x-axis) when a purified antibody binds its antigen (VEGF) immobilized on the CM5 biochip at 35°C. K_d values were calculated by the ratio of K_{on} to K_{off}. A ratio of a variant K_d to that of hAB1 for each variant antibody was also calculated. hAB1 contains SEQ1 and SEQ55 (which, if in the form of Fab, would be essentially Fab-12 as described in Chen et al. (1999) J. Mol. Biol. 293:865-881); hAB7 contains SEQ1 and SEQ101; hAB8 contains SEQ1 and SEQ100; hAB13 contains SEQ1 and SEQ56 (which, if in the form of Fab, would be Y0317 as described in Chen et al. (1999) J. Mol. Biol. 293:865-881); hAB14 contains SEQ36 and SEQ100; hAB15 contains SEQ26 and SEQ100; hAB16 contains SEQ28 and SEQ100; hAB35 contains SEQ28 and SEQ106; hAB36 contains SEQ28 and SEQ107; hAB37 contains SEQ28 and SEQ108; hAB38 contains SEQ28 and SEQ109; hAB39 contains SEQ28 and SEQ110, as V_L and V_H, respectively.

Figure 3 shows the affinity analysis of six humanized anti-VEGF antibodies using BIAcore biosensor. The antibodies were incubated in 1xPBS buffer at 4°C, 37°C, or 42°C for 16 hours before the assay. The measurement was done by measuring the change of SPR units (y-axis) vs time (x-axis) when a purified antibody binds its antigen (VEGF) immobilized on the CM5 biochip at 25°C. Both the on-rate and off-rate changes were determined from the data fitting using 1:1 Langmuir binding model, whereas K_ds were determined by the ratio of K_{off} to K_{on}. Composition of the V_L and V_H, respectively, is listed as following: hAB1 contains SEQ1 and SEQ55; hAB2 contains SEQ1 and SEQ70; hAB3 contains SEQ1 and SEQ67; hAB7 contains SEQ1 and SEQ101; hAB8 contains SEQ1 and SEQ100; hAB13 contains SEQ1 and SEQ56, as V_L and V_H, respectively.

Figure 4 summarized the data of the stability of the humanized anti-VEGF antibodies shown in Figure 3. The y-axis shows the percentage of the antibody remain active in binding to the immobilized VEGF antigen using BIAcore at 25°C after the purified antibody is incubated at 4°C, 37°C and 42°C for 16 hours as described in Figure 3. The maximal binding of each antibody at different conditions was expressed as percentile of that of 4°C.

Figure 5 shows the expression of the humanized anti-VEGF antibodies using E.coli expression system as described below. The expression level of each antibody fragment was evaluated by applying the same volume of the purified material from the same fraction as shown detected by SDS-PAGE/Coomassie blue staining. Composition of the V_L and V_H, respectively, is listed as following: hAB1 contains SEQ1 and SEQ55; hAB35 contains SEQ28 and SEQ106; hAB36 contains SEQ28 and SEQ107; hAB37 contains SEQ28 and SEQ108; hAB38 contains SEQ28 and SEQ109; hAB39 contains SEQ28 and SEQ110, as V_L and V_H, respectively.

Figure 6 shows expression of humanized anti-VEGF antibody in a eukaryotic system as described below. The expression level of each antibody fragment was evaluated by applying the same volume of the purified material from the same fraction as shown detected by SDS-PAGE/Coomassie blue staining. Composition of the VL and VH, respectively, is listed as following: hAB1 contains SEQ1 and SEQ55; hAB2 contains SEQ1 and SEQ70; hAB3 contains SEQ1 and SEQ67; hAB5 contains SEQ1 and SEQ75; hAB7 contains SEQ1 and SEQ101; hAB8 contains SEQ1 and SEQ100; hAB13 contains SEQ1 and SEQ56, as V_L and V_H, respectively.

Figure 7 illustrates a map of the vector used for expression of soluble antibody fragments in *E. coli*.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides an innovative methodology for engineering proteins, in particular antibodies, with desired structures and functions. In one aspect, methods are provided for designing and selecting antibodies against human antigens with high affinity and specificity. In some particular embodiments, methods are provided for designing and selecting humanized or fully human antibodies against vascular endothelial growth factor (VEGF) with high affinity and specificity. In another aspect, compositions, kits and methods are provided for using these antibodies and their derivatives to inhibit angiogenesis in vitro, and for diagnosing or treating diseases associated with abnormal angiogenesis such as cancer, rheumatoid arthritis, ischemic-reperfusion related brain edema and injury, cortical ischemia, ovarian hyperplasia and hypervascularity, endometriosis, psoriasis, diabetic retinopaphy, and other ocular angiogenic diseases.

1. VEGF and Antibodies Against VEGF

VEGF is a key angiogenic factor in development and is involved in the growth of solid tumor by stimulating endothelial cells. A murine monoclonal antibody was found to block VEGF-dependent cell proliferation and slow the tumor growth in vivo (Kim KJ, Li B, Winer J, Armanini M, Gillett N, Phillips HS, Ferrara N (1993) Nature 362, 841-844). This murine antibody was humanized (Presta LG, Chen H, O'Connor SJ, Chisholm V, Meng YG, Krummen L, Winkler M, Ferrara N (1997) Cancer Res. 57, 4593-4599; Baca M, Presta LG, O'Connor SJ, Wells JA (1997) J Biol Chem 272, 10678-10684) and affinity-matured by using phage-display and off-rate selection (Chen Y, Wiesmann C, Fuh G, Li B, Christinger HW, McKay P, de Vos AM (1999) J Mol Biol 293, 865-881). X-ray structure for the complex formed between VEGF and the parental antibody was reported (Muller YA, Chen Y, Christinger HW, Li B, Cunningham, BC, Lowman HB, de Vos AM (1998) Structure 6, 1153-1167.), as well as the one formed between VEGF and the matured antibody (Chen Y, Wiesmann C, Fuh G, Li B, Christinger HW, McKay P, de Vos AM, Lowman HB (1999) J. Mol Biol 293, 865-881). In addition, US Patent application serial No: 09/056,160, publication No: 2002/0032315, discloses certain anti-VEGF antibodies. These publications on VEGF and anti-VEGF antibodies are herein incorporated by reference in their entirety.

The present invention provides novel anti-VEGF antibodies that bind to human VEGF with high affinity and specificity. These anti-VEGF antibodies are humanized and optimized with some important attributes in binding affinity, stability, expression efficiency which are desirable for research, diagnostic and therapeutic applications.

The binding affinity of the selected anti-VEGF antibodies of the present invention to a human VEGF, represented by dissociation constant K_d, is optionally lower than 100 nM, optionally lower than 10 nM, optionally lower than 8 nM, optionally lower than 8 nM, optionally lower than 5 nM, optionally lower than 1 nM, optionally lower than 0.8 nM, optionally lower than 0.5 nM, optionally lower than 0.2 nM, optionally lower than 0.1 nM, optionally lower than 0.08 nM, optionally lower than 0.05 nM, optionally lower than 0.01 nM, or optionally lower than 0.005 nM, in the form of scFv, Fab, or other form of antibody measured at a temperature of about 4°C, 25°C, 37°C or 42°C.

Preferably, K_d of the selected anti-VEGF antibodies of the present invention, if in the form of scFv and measured at a temperature of 35-37°C, is lower than 1 nM, optionally lower than 0.8 nM, optionally lower than 0.5 nM, optionally lower than 0.2 nM, optionally lower than 0.1 nM, optionally lower than 0.08 nM, optionally lower than 0.05 nM, optionally lower than 0.01 nM, or optionally lower than 0.005 nM.

Also preferably, K_d of the selected anti-VEGF antibodies of the present invention, if in the form of Fab and measured at a temperature of 35-37°C, is lower than 0.5 nM, optionally lower than 0.2 nM, optionally lower than 0.1 nM, optionally lower than 0.08 nM, optionally lower than 0.05 nM, optionally lower than 0.01 nM, or optionally lower than 0.005 nM.

In inhibition of VEGF-induced proliferation of endothelial cells in vitro the selected anti-VEGF antibodies of the present invention preferably has an ED_{50} (effective dose for inhibition of 50% cell proliferation) equal to or lower than 10 nM, optionally lower than 5 nM, optionally lower than 1 nM, optionally lower than 0.5 nM, optionally lower than 0.1 nM, optionally lower than 0.05 nM, or optionally lower than 0.01 nM in the form of scFv, Fab or other form of antibody.

In a preferred embodiment, an anti-VEGF antibody is provided that has a light chain variable region comprising the amino acid sequence of:

$X_1X_2X_3X_4TQX_5PSX_6X_7SX_8X_9X_{10}GX_{11}X_{12}X_{13}X_{14}IX_{15}C\underline{X_{16}}\underline{X_{17}}\underline{SX_{18}}\underline{X_{19}}\underline{IX_{20}}\underline{X_{21}}\underline{X_{22}}\underline{X_{23}}\underline{X_{24}}$
$WYQQX_{25}PGX_{26}APX_{27}X_{28}LX_{29}Y\underline{X_{30}}\underline{X_{31}}\underline{X_{32}}\underline{X_{33}}\underline{LX_{34}}\underline{X_{35}}GVX_{36}X_{37}RFSGX_{38}X_{39}SGTDF$
$X_{40}LTIX_{41}X_{42}LQX_{43}X_{44}DX_{45}AX_{46}YYC\underline{QQX_{47}}\underline{X_{48}}\underline{X_{49}}\underline{X_{50}}\underline{PX_{51}}\underline{T}FGX_{52}GTKX_{53}X_{54}IK$,

wherein the underlined regions are designated as V_L/CDR1, V_L/CDR2, and V_L/CDR3, respectively, whereas the rest of the region is designated as framework, and wherein the position designated as "X" could be amino acids listed below:

X_1 : D, E or A
X_2 : I, or T
X_3 : V, E, K, R, Q, or T
X_4 : M, or L
X_5 : S, or T
X_6 : S, or T
X_7 : L, or V
X_8 : A, or V
X_9 : S, or T
X_{10} : P, V, L, A, or I
X_{11} : E, or D
X_{12} : R, or T
X_{13} : A, or V I
X_{14} : T, or A
X_{15} : T, S, or A
X_{16} : S, R, N, K, H, or Q
X_{17} : A, or S
X_{18} : Q, or R
X_{19} : S, D, A, or P
X_{20} : S, G, R, T, or Y
X_{21} : T, N, S, D, or K
X_{22} : Y, or D
X_{23} : L, or I
X_{24} : A, N, or T
X_{25} : K, or I
X_{26} : Q, K, T, or I
X_{27} : R, K, Q, N, H, S, or E
X_{28} : V, or L
X_{29} : I, or V
X_{30} : F, A, G , D, or S
X_{31} : A , or T
X_{32} : S, or T
X_{33} : N, S, R, or T
X_{34} : A, H, or Q
X_{35} : S , or G
X_{36} : P, T
X_{37} : S, N, D, G, or Y

X_{38} : S, or T
X_{39} : G, or R
X_{40} : T, or A
X_{41} : S, or R
X_{42} : S, or R
X_{43} : P, or A
X_{44} : E , or D
X_{45} : F, V, or S
X_{46} : V, T, I , A , or S
X_{47} : Y, or S
X_{48} : S, Y, or N
X_{49} : S , or T
X_{50} : T, V, A, P, K, G, S, or I
X_{51} : W , or Y
X_{52} : Q , or G
X_{53} : V, or L
X_{54}: E, D , or A.

Such preferred light chain variable domain sequences may be combined with the preferred heavy chain variable domain sequences, or with other heavy chain variable domain sequences, provided that the antibody so produced binds to human VEGF with a desired affinity.

In another preferred embodiment, an anti-VEGF antibody is provided that has a light chain variable region comprising the amino acid sequence of:

$X_1X_2X_3$LTQPPSX_4SX_5TPGQX_6VTISCS$\underline{GX_7X_8SNX_9GX_{10}NX_{11}VX_{12}}$WYQQ$X_{13}PGX_{14}$A PK$X_{15}LX_{16}Y\underline{X_{17}NX_{18}X_{19}RPS}GVPX_{20}RX_{21}SGSX_{22}SX_{23}$TSASLAISGL$X_{24}$SEDEADYY C$\underline{X_{25}X_{26}WDDSLX_{27}GYV}FGX_{28}GX_{29}$LTVL, wherein the underlined regions are designated as V_L/CDR1, V_L/CDR2, and V_L/CDR3, respectively, whereas the rest of the region is designated as framework, and wherein the position designated as "X" could be amino acids listed below:

X_1 : Q L, or N
X_2 : P A F, or S
X_3 : V, or M
X_4 : A, or T
X_5 : G, or A
X_6 : R, or S
X_7 : S, or T
X_8 : S, T Y, or N
X_9 : I , or V
X_{10} : S, or R
X_{11} : S, P, N, A, or T

X_{12} : N, T , or Y
X_{13} : L, or F
X_{14} : T, or A
X_{15} : V, L, or F
X_{16} : M, or I
X_{17} : G, T, or S
X_{18} : N, or D
X_{19} : Q, or E
X_{20} : D, or E
X_{21} : F, or L
X_{22} : K, or R
X_{23} : G, or A
X_{24} : Q, L, or R
X_{25} : A, or G
X_{26} : A, S, or T
X_{27} : N, S, or T
X_{28} : T, or A
X_{29} : K, or Q.

Such preferred light chain variable domain sequences may be combined with the preferred heavy chain variable domain sequences, or with other heavy chain variable domain sequences, provided that the antibody so produced binds to human VEGF with desired affinity.

In yet another preferred embodiment, an anti-VEGF antibody is provided that has a light chain variable region comprising the amino acid sequence of: QSALTQPPSVSGAPGQRVTISCTGRSSNIGAGHDVHWYQQLPGTAPKLLIYANDQ RPSGVPDRFSDSKSGTSASLGISGLRSEDEADYFCATWDDSLHGYVFGTGTKVTV L (SEQ ID NO:54), such preferred light chain variable domain sequences may be combined with the heavy chain variable domain sequences, or with other heavy chain variable domain sequences, provided that the antibody so produced binds to human VEGF with desired affinity.

In yet another preferred embodiment, an anti-VEGF antibody is provided that has a heavy chain variable region comprising the amino acid sequence of: X_1X_2QLVX_3SGGGX_4VQPGGX_5LRLX_6CAX_7S$\underline{GX_8X_9X_{10}X_{11}X_{12}X_{13}GX_{14}N}WX_{15}$RQAP GKG$X_{16}$EWVG$\underline{WX_{17}NTX_{18}X_{19}GX_{20}X_{21}TYX_{22}X_{23}X_{24}FX_{25}R}RX_{26}TX_{27}SX_{28}X_{29}X_{30}SKX_{31}X_{32}X_{33}YLQX_{34}$NSLRAEDTAVYYCA$\underline{X_{35}YPX_{36}YYGX_{37}SHWYFDV}WX_{38}$QGTLVTV SS, wherein the underlined regions are designated as CDR1, CDR2, and CDR3,

respectively, whereas the rest of the region is designated as framework according to Kabat nomenclature, and wherein the position designated as "X" could be amino acids listed below:

X_1 : E, or Q
X_2 : V, or G
X_3 : Q, or E
X_4 : V, or L
X_5 : S, or T
X_6 : S T, or R
X_7 : A, or V
X_8 : Y, or F
X_9 : T, D, N, S, or A
X_{10} : F, or L
X_{11} : T, D, Y, A, S, or N
X_{12} : N, H, or S
X_{13} : Y, or F
X_{14} : M, L, I, or V
X_{15} : I, V, or L
X_{16} : L, or P
X_{17} : I, or V
X_{18} : Y, or N
X_{19} : T, or N
X_{20} : E, or A
X_{21} : P, T, or S
X_{22} : A, or V
X_{23} : A, H, Q, P, D, or E
X_{24} : D, or E
X_{25} : K, or T
X_{26} : V, F, or L
X_{27} : F, or I
X_{28} : L, or R
X_{29} : D, or N
X_{30} : T, or N
X_{31} : S, or N
X_{32} : T, Q, P, or K
X_{33} : A, V, or P
X_{34} : L, or M
X_{35} : K, or R
X_{36} : H, or Y
X_{37} : S, R, or T
X_{38} : G, or A.

In yet another preferred embodiment, an anti-VEGF antibody is provided that has a heavy chain variable region comprising the amino acid sequence of:

$X_1X_2QLVX_3SGGGX_4VQPGGX_5LRLX_6CAX_7SGX_8X_9X_{10}X_{11}X_{12}X_{13}GX_{14}NWX_{15}RQAP$
$GKGX_{16}EWVGWX_{17}NTX_{18}X_{19}GX_{20}X_{21}TYX_{22}X_{23}X_{24}FX_{25}RRX_{26}TX_{27}SX_{28}X_{29}X_{30}SKX_3$
$_1X_{32}X_{33}YLQX_{34}NSLRAEDTAVYYCAX_{35}X_{36}X_{37}X_{38}X_{39}X_{40}X_{41}X_{42}X_{43}X_{44}X_{45}YX_{46}DX_{47}$
$WX_{48}QGTLVTV$, wherein the underlined regions are designated as CDR1, CDR2, and CDR3, respectively, whereas the rest of the region is designated as framework according to Kabat nomenclature, and wherein the position designated as "X" could be amino acids listed below:

X_1 : E, or Q
X_2 : V, or G
X_3 : Q, or E
X_4 : V, or L
X_5 : S, or T
X_6 : S T, or R
X_7 : A, or V
X_8 : Y, or F
X_9 : T, D, N, S, or A
X_{10} : F, or L
X_{11} : T, D, Y, A, S, or N
X_{12} : N, H, or S
X_{13} : Y, or F
X_{14} : M, L, I, or V
X_{15} : I, V, or L
X_{16} : L, or P
X_{17} : I, or V
X_{18} : Y, or N
X_{19} : T, or N
X_{20} : E, or A
X_{21} : P, T, or S
X_{22} : A, or V
X_{23} : A, H, Q, P, D, or E
X_{24} : D, or E
X_{25} : K, or T
X_{26} : V, F, or L
X_{27} : F, or I
X_{28} : L, or R
X_{29} : D, or N
X_{30} : T, or N
X_{31} : S, or N
X_{32} : T, Q, P, or K

X_{33} : A, V, or P
X_{34} : L, or M
X_{35} : K R, or H
X_{36} : Y A D, or S
X_{37}: P R S, or G
X_{38} : Y H, or D
X_{39} : Y, or F
X_{40} : Y N S, or H
X_{41} : G , or S
X_{42} : S, T, R, G, or A
X_{43} : S, Y, C, or T
X_{44} : H, P, C, N, Q, or S
X_{45} : W, Q, or C
X_{46} : F , or L
X_{47} : V, L, or Y
X_{48} : G, or A.

Such preferred heavy chain variable domain sequences may be combined with the preferred light chain variable domain sequences or with other light chain variable domain sequences, provided that the antibody so produced binds to human VEGF with desired affinity.

In one embodiment, the invention provides an anti-VEGF antibody that preferably contains a light chain variable domain comprising the amino acid sequences of one of the following: SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, and SEQ ID NO:53, further preferably SEQ ID NO:14, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:44, SEQ ID NO:47, or SEQ ID NO:54. Such preferred light chain variable domain sequences may be combined with the preferred heavy chain variable

domain sequences or with other heavy chain variable domain sequences, provided that the antibody so produced binds to human VEGF with desired affinity.

In another embodiment, the invention provides an anti-VEGF antibody that preferably contains a heavy chain variable domain comprising the amino acid sequence of one of the following: SEQ ID NO: ID NO:57, SEQ ID NO: ID NO:58, SEQ ID NO: ID NO:59, SEQ ID NO: ID NO:60, SEQ ID NO: ID NO:65, SEQ ID NO: ID NO:66, SEQ ID NO: ID NO:69, SEQ ID NO: ID NO:71, SEQ ID NO: ID NO:72, SEQ ID NO: ID NO:73, SEQ ID NO: ID NO:74, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:80, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86, SEQ ID NO:87, further preferably SEQ ID NO:61, or SEQ ID NO:62, or SEQ ID NO:63, or SEQ ID NO:64, or SEQ ID NO:67, or SEQ ID NO:68, or SEQ ID NO:70, or SEQ ID NO:75, or SEQ ID NO:83, SEQ ID NO:88, or SEQ ID NO:89, or SEQ ID NO:90, or SEQ ID NO:91, or SEQ ID NO:92, or SEQ ID NO:93, or SEQ ID NO:94, or SEQ ID NO:95, SEQ ID NO:96, SEQ ID NO:97, SEQ ID NO:98, or SEQ ID NO:99, SEQ ID NO:100, or SEQ ID NO:101, or SEQ ID NO:102, or SEQ ID NO:103, or SEQ ID NO:104, or SEQ ID NO:105, or SEQ ID NO:106, SEQ ID NO:107, or SEQ ID NO:108, or SEQ ID NO:109, SEQ ID NO:110; or one of SEQ ID NOs:285-310. Such preferred heavy chain variable region sequence may be combined with the preferred light chain variable region sequence or with other light chain variable region sequence, provided that the antibody so produced binds to human VEGF with desired affinity.

In yet another embodiment, the invention provides an anti-VEGF antibody that preferably contains the CDR1 regions of the light chain variable domain comprising the amino acid sequence of one of the following: SEQ ID NO:164, SEQ ID NO:165, SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, SEQ ID NO:171, SEQ ID NO:172, SEQ ID NO:173, SEQ ID NO:174, SEQ ID NO:175, SEQ ID NO:176, SEQ ID NO:177, SEQ ID NO:178, SEQ ID NO:179, SEQ ID NO:180, SEQ ID NO:181, SEQ ID NO:182, SEQ ID NO:183, SEQ ID NO:184, SEQ ID NO:185, SEQ ID NO:186, SEQ ID NO:187, SEQ ID NO:188, SEQ ID NO:189, SEQ ID NO:190, SEQ ID NO:191, SEQ ID NO:192, SEQ ID NO:193, and SEQ ID NO:194. Such preferred CDR1 sequence of light chain variable domain may be combined with other regions of preferred

light chain, or of other light chain, and the preferred heavy chain variable region sequence or with other heavy chain variable region sequence, provided that the antibody so produced binds to human VEGF with desired affinity.

In yet another embodiment, the invention provides an anti-VEGF antibody that preferably contains the CDR2 regions of the light chain variable domain comprising the amino acid sequence of one of the following: SEQ ID NO:195, SEQ ID NO:196, SEQ ID NO:197, SEQ ID NO:198, SEQ ID NO:199, SEQ ID NO:200, SEQ ID NO:201, SEQ ID NO:202, SEQ ID NO:203, SEQ ID NO:204, SEQ ID NO:205, SEQ ID NO:206, SEQ ID NO:207, SEQ ID NO:208, and SEQ ID NO:209. Such preferred sequence of light chain variable domain may be combined with other regions of preferred light chain, or of other light chain, and the preferred heavy chain variable region sequence or with other heavy chain variable region sequence, provided that the antibody so produced binds human VEGF.

In yet another embodiment, the invention provides an anti-VEGF antibody that preferably contains the CDR3 regions of the light chain variable domain comprising the amino acid sequence of one of the following: SEQ ID NO:210, SEQ ID NO:211, SEQ ID NO:212, SEQ ID NO:213, SEQ ID NO:214, SEQ ID NO:215, SEQ ID NO:216, SEQ ID NO:217, SEQ ID NO:218, SEQ ID NO:219, SEQ ID NO:220, SEQ ID NO:221, SEQ ID NO:222, SEQ ID NO:223, SEQ ID NO:224, SEQ ID NO:225, SEQ ID NO:226, SEQ ID NO:227, and SEQ ID NO:228. Such preferred CDR3 sequence of light chain variable domain may be combined with other regions of preferred light chain, or of other light chain, and the preferred heavy chain variable region sequence or with other heavy chain variable region sequence, provided that the antibody so produced binds to human VEGF with desired affinity.

In yet another embodiment, the invention provides an anti-VEGF antibody that preferably contains the framework regions of the light chain variable domain comprising the amino acid sequence of one of the following: SEQ ID NO:229, SEQ ID NO:230, SEQ ID NO:231SEQ ID NO:233, SEQ ID NO:234, SEQ ID NO:236, SEQ ID NO:238, SEQ ID NO:239, SEQ ID NO:240, SEQ ID NO:241, SEQ ID NO:242, SEQ ID NO:243, SEQ ID NO:244, SEQ ID NO:245, SEQ ID NO:246, SEQ ID NO:247, SEQ ID NO:248, SEQ ID NO:249, SEQ ID NO:250, SEQ ID NO:252, SEQ ID NO:253, SEQ ID NO:254,

SEQ ID NO:256, SEQ ID NO:257, SEQ ID NO:258, SEQ ID NO:259, SEQ ID NO:260, SEQ ID NO:261, SEQ ID NO:262, SEQ ID NO:264, SEQ ID NO:266, SEQ ID NO:267, SEQ ID NO:268, and SEQ ID NO:269; further preferably contains the framework regions of the light chain variable domain comprising the amino acid sequence of one of the following: SEQ ID NO:232, SEQ ID NO:235, SEQ ID NO:237, SEQ ID NO:251, SEQ ID NO:255, SEQ ID NO:263, and SEQ ID NO:265. Such preferred framework region sequence of light chain variable region may be combined with CDR regions of preferred light chain, or of other light chain, and the preferred heavy chain variable region sequence or with other heavy chain variable region sequence, provided that the antibody so produced binds to human VEGF with desired affinity.

In one embodiment, the invention provides an anti-VEGF antibody that preferably contains the CDR1 regions of the heavy chain variable domain comprising the amino acid sequence of one of the following: $GX_1X_2X_3X_4X_5X_6GX_7N$, wherein the position designated as "X" could be amino acids listed below:

X_1: Y, or F
X_2: D, N, T, S, or A
X_3: F, or L
X_4: T, D, S, Y, A, or N
X_5: H, N, or S
X_6: Y, or F
X_7: M, L, I , or V.

Further preferably the CDR1 region of the heavy chain variable domain comprises the amino acid sequence of one of the following: SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:114, SEQ ID NO:115, SEQ ID NO:116, SEQ ID NO:117, SEQ ID NO:118, SEQ ID NO:119, SEQ ID NO:120, SEQ ID NO:121, SEQ ID NO:122, SEQ ID NO:123, SEQ ID NO:124, SEQ ID NO:125, SEQ ID NO:126, SEQ ID NO:127, SEQ ID NO:128, SEQ ID NO:129, SEQ ID NO:130, SEQ ID NO:131, SEQ ID NO:132, SEQ ID NO:133, SEQ ID NO:134, and SEQ ID NO:135. Such preferred CDR1 sequence of heavy chain variable domain may be combined with other regions of preferred heavy chain or of other heavy chain, and the preferred light chain variable region sequence or with other light chain variable region sequence, provided that the antibody so produced binds human VEGF.

In one embodiment, the invention provides an anti-VEGF antibody that preferably contains CDR2 of the heavy chain variable domain comprising the amino acid sequence of one of the following: $WX_1NTX_2X_3GEX_4TYX_5X_6X_7FX_8R$, wherein the position designated as "X" could be amino acids listed below:

X_1: I, or V
X_2: Y, or N
X_3: T, or N
X_4: P, T, or S
X_5: A, or V
X_6: A, Q, P, H, D, or E
X_7: D, or E
X_8: K, or T

Further preferably CDR2 of the heavy chain variable domain comprise the amino acid sequence of one of the following: SEQ ID NO:136,SEQ ID NO:137, SEQ ID NO:138, SEQ ID NO:139, SEQ ID NO:140, SEQ ID NO:141, SEQ ID NO:142, SEQ ID NO:143, SEQ ID NO:144, SEQ ID NO:145, SEQ ID NO:146, SEQ ID NO:147, SEQ ID NO:148, SEQ ID NO:149, SEQ ID NO:150, SEQ ID NO:151, SEQ ID NO:152, SEQ ID NO:153, SEQ ID NO:154, SEQ ID NO:155, and SEQ ID NO:156. Such preferred CDR2 sequence of heavy chain variable domain may be combined with other regions of preferred heavy chain or of other heavy chain, and the preferred light chain variable region sequence or with other light chain variable region sequence, provided that the antibody so produced binds to human VEGF with desired affinity.

In one embodiment, the invention provides an anti-VEGF antibody that contains CDR3 of the heavy chain variable domain comprising the amino acid sequence: $KYPX_1YYGX_2SHWYFDV$, wherein the position designated as "X" could be amino acids listed below: X_1: Y, or H, and X_2: R.

Preferably, the anti-VEGF antibody has CDR3 of the heavy chain variable domain comprising the amino acid sequence of one of SEQ ID NOs:311-337, and the following sequences:

HSRHYYGSSPQYFDV
KYGYYYGSSHWYFDV
KYPHYYGASHWYFDV
KYPHYYGGCHWYFDV
KYPHYYGGSHWYFDV
KYPHYYGGYNQYFDV

KYPHYYGRSHWYFDV
KYPHYYGRSQWYLDV
KYPHYYSRTCQYFDV
KYPHYYSSSHWYFDV
KYPYFYGSSHWYFDV
KYPYYHGSSHWYFDV
KYPYYNGSSHWYFDV
KYPYYNSTSHWYFDV
KYPYYSGTSHWYFDV
KYPYYSGTSHWYFDY
KYPYYYGRSHWYFDV
KYPYYYGSSHWYFDV
KYPYYYGSSSWYFDV
KYPYYYSTSHWYFDV
KYRDFNGSSHWYFDV
KYSYYYGSSHWYFDV
RARHYYGSSHCYFDL
RDSHYYGSSHQYFDL
KYPHYYGTSHWYFDV
KYPHYYGSSHWYFDV
KYPYYYGTSHWYFDV.

Such preferred CDR3 sequence of heavy chain variable domain may be combined with other regions of preferred heavy chain or of other heavy chain, and the preferred light chain variable region sequences or with other light chain variable domain sequences, provided that the antibody so produced binds to human VEGF with desired affinity.

In one embodiment, the invention provides an anti-VEGF antibody that preferably contains the framework region of the heavy chain variable domain comprising the amino acid sequences of one of the following:

X_1VQLVX_2SGGGVX_3VQPGGX_4LRLSX_5CAX_6S/CDR1/WX_7RQAPGKGLEWVG/CDR2/RX_8TFTISRDNSKNTLYLQMNSLRAEDTAVYYCA/CDR3/WGQGTLVTVSS, wherein the position designated as "X" could be amino acids listed below:

X_1 : E, or Q
X_2 : Q, or E
X_3 : V, or L
X_4 : S, or T
X_5 : S, T, or R
X_6 : A, or V
X_7 : I, or V
X_8 : F, or V

X_9 : F, or I
X_{10} : L, or R
X_{11} : T, or N
X_{12} : S, or N
X_{13} : T, Q, or K
X_{14} : A, or V
X_{15} : M, or L
X_{16} : G, or A.

In one embodiment, the invention provides an anti-VEGF antibody that preferably contains a light chain variable domain and a heavy chain variable domain comprising the amino acid sequences of one of the following V_L and V_H pairs: SEQ ID NO:1 and SEQ ID NO:70; SEQ ID NO:1 and SEQ ID NO:67; SEQ ID NO:1 and SEQ ID NO:75; SEQ ID NO:1 and SEQ ID NO:83; SEQ ID NO:14 and SEQ ID NO:55; SEQ ID NO:1 and SEQ ID NO:101; SEQ ID NO:1 and SEQ ID NO:100; SEQ ID NO:14 and SEQ ID NO:102; SEQ ID NO:1 and SEQ ID NO:103; SEQ ID NO:1 and SEQ ID NO:104; SEQ ID NO:1 and SEQ ID NO:105; SEQ ID NO:36 and SEQ ID NO:100; SEQ ID NO:26 and SEQ ID NO:100; SEQ ID NO:28 and SEQ ID NO:100; SEQ ID NO:37 and SEQ ID NO:100; SEQ ID NO:44 and SEQ ID NO:100; SEQ ID NO:54 and SEQ ID NO:100; SEQ ID NO:47 and SEQ ID NO:100, further preferably SEQ ID NO:28 and SEQ ID NO:61; SEQ ID NO:28 and SEQ ID NO:62; SEQ ID NO:28 and SEQ ID NO:63; SEQ ID NO:28 and SEQ ID NO:64; SEQ ID NO:28 and SEQ ID NO:68; SEQ ID NO:28 and SEQ ID NO:85; SEQ ID NO:28 and SEQ ID NO:86; SEQ ID NO:28 and SEQ ID NO:87; SEQ ID NO:28 and SEQ ID NO:88; SEQ ID NO:28 and SEQ ID NO:89; SEQ ID NO:28 and SEQ ID NO:90; SEQ ID NO:28 and SEQ ID NO:91; SEQ ID NO:28 and SEQ ID NO:92; SEQ ID NO:28 and SEQ ID NO:93; SEQ ID NO:28 and SEQ ID NO:94; SEQ ID NO:28 and SEQ ID NO:95; SEQ ID NO:28 and SEQ ID NO:96; SEQ ID NO:28 and SEQ ID NO:97; SEQ ID NO:28 and SEQ ID NO:98; SEQ ID NO:28 and SEQ ID NO:99; SEQ ID NO:28 and SEQ ID NO:106; SEQ ID NO:28 and SEQ ID NO:107; SEQ ID NO:28 and SEQ ID NO:108; SEQ ID NO:28 and SEQ ID NO:109; and SEQ ID NO:28 and SEQ ID NO:110.

The antibodies resulted from combination of the full-length V_H and V_L, V_H/CDR, V_H/FR, V_L/CDR, V_L/FR (e.g., the hit variants provided in Figures 1A and 1B and amino acid sequences shown in Figures 1C and 1D) disclosed in the present invention are within

the scope of the present invention; and such combination does not include the anti-VEGF antibodies disclosed in US Patent application serial No: 09/056,160, publication No: 2002/0032315.

2. Methodology for Designing and Constructing Humanized or Fully Human Anti-VEGF Antibodies

The antibodies of the present invention are designed using innovative methods involving construction and selection of protein libraries *in silico* and *in vitro*. The following describes some aspects of the methodology. More detailed description of the methodology appears in U.S. Patent Application Nos: 10/443,134, 10/153,159, 10/153,176, 10/125,687, and 60/284,407, which are incorporated herein by reference in their entirety.

According to the present invention, an innovative methodology is provided for efficiently generating and screening protein libraries for optimized proteins with desirable biological functions, such as improved binding affinity towards biologically and therapeutically important target molecules. The methodology is used to optimize proteins by generating novel variants of a protein with enhanced properties. In particular, this methodology is used to design libraries for humanization of non-human antibodies and to optimize the affinity and other attributes of antibodies. Novel variants of amino acids and nucleic acids of antibodies are generated with human or human-like sequences while their binding affinity, stability, and expression efficiency are improved significantly.

The inventive process is carried out computationally in a high throughput manner by mining the ever-expanding databases of protein sequences of all organisms, especially human and by relating their specific sequences or their variants with functional enhancement such as binding affinity and stability that are tested experimentally. By using the inventive methodology, an expanded and yet functionally biased library of proteins such as antibodies can be constructed based on computational evaluation of extremely diverse protein sequences and functionally relevant structures in silico and subsequently tested by experimental screening and selection in vitro or in vivo.

In one aspect of the invention, a method is provided for designing and selecting protein(s) with desirable function(s). The method is preferably implemented in a computer through *in silico* selection of protein sequences based on the amino acid sequence of a target structural/functional motif or domain in a lead protein, herein after referred to as the "lead sequence". The lead sequence is employed to search databases of protein sequences. The choice of the database depends on the specific functional requirement of the designed motifs. For example, if the lead protein is an enzyme and the target motif includes the active site of the enzyme, databases of proteins/peptides of a particular origin, organism, species or combinations thereof, may be queried using various search criteria to yield a hit list of sequences each of which can substitute the target motif in the lead protein. A similar approach may be used for designing other motifs or domains of the lead protein. The designed sequences for each individual motif/domain may be combined to generate a library of designed proteins. In addition, to reduce immunogenicity of the designed proteins for human applications such as therapeutics or diagnosis, databases of proteins of human origin or humanized proteins are preferably searched to yield the hit list of sequences, especially for motifs derived from sites of the lead protein that serves as the scaffolding of the lead proteins suc as the frameworks of an antibody. The library of designed proteins can be tested experimentally to yield proteins with improved biological function(s) over the lead protein.

In a particular aspect of the invention, the inventive methodology is implemented in designing antibodies that are diverse in sequence and yet functionally related to each other. Based on the designed antibody sequences, a library of antibodies can be constructed to include diverse sequences in the complementary determining regions (CDRs) and/or humanized frameworks (FRs) of a non-human antibody in a high throughput manner. This library of antibodies can be screened against a wide variety of target molecules for novel or improved functions.

In yet another aspect of the invention, a method is provided for in silico selection of antibody sequences based on the amino acid sequence of a region in a lead antibody, herein after referred to as the "lead sequence". The lead sequence is employed to search databases of protein sequences. The choice of the database depends on the specific

functional requirement of the designed motifs. For example: in order to design the framework regions of variable chains for therapeutic application, collections of protein sequences that are evolutionarily related such as fully human immunoglobulin sequences and human germline immunoglobulin sequences should be used except for a few structurally critical sites. This would reduce the immunogenic response by preserving the origin of the sequences by introducing as few foreign mutants as possible in this highly conserved region (for framework regions). On the other hand, diverse sequence databases such as immunoglobulin sequences of various species or even unrelated sequence in genbank can be used to design the CDRs in order to improve binding affinity with antigens in this highly variable region. By using the method, a library of diverse antibody sequences can be constructed and screened experimentally in vitro or in vivo for antibody mutants with improved or desired function(s).

In one embodiment, the method comprises the steps of: providing an amino acid sequence of the variable region of the heavy chain (V_H) or light chain (V_L) of a lead antibody; identifying the amino acid sequences in the CDRs of the lead antibody; selecting one of the CDRs in the V_H or V_L region of the lead antibody; providing an amino acid sequence that comprises at least 3 consecutive amino acid residues in the selected CDR, the selected amino acid sequence being a lead sequence; comparing the lead sequence with a plurality of tester protein sequences; and selecting from the plurality of tester protein sequences at least two peptide segments that have at least 15% sequence identity with the lead sequence, the selected peptide segments forming a hit library.

The method may further comprise the step of: constructing a nucleic acid library comprising DNA segments encoding the amino acid sequences of the hit library.

Optionally, the method may further comprise the steps of: building an amino acid positional variant profile of the hit library; converting amino acid positional variant profile of the hit library into a nucleic acid positional variant profile by back-translating the amino acid positional variants into their corresponding genetic codons; and constructing a degenerate nucleic acid library of DNA segments by combinatorially combining the nucleic acid positional variants.

Optionally, the genetic codons may be the ones that are preferred for expression in bacteria. Optionally, genetic codons may be the ones that can reduce the size chosen

such that the diversity of the degenerate nucleic acid library of DNA segments is within the experimentally coverable diversity without undue experimental efforts, for example, to be below $1x10^7$, and preferably below $1x10^6$.

In another embodiment, the method comprises the steps of: providing an amino acid sequence of the variable region of the heavy chain (V_H) or light chain (V_L) of a lead antibody; identifying the amino acid sequences in the CDRs and FRs of the lead antibody; selecting one of the CDRs in the V_H or V_L region of the lead antibody; providing a first amino acid sequence that comprises at least 3 consecutive amino acid residues in the selected CDR, the selected amino acid sequence being a CDR lead sequence; comparing the CDR lead sequence with a plurality of CDR tester protein sequences; selecting from the plurality of CDR tester protein sequences at least two peptide segments that have at least 15% sequence identity with the CDR lead sequence, the selected peptide segments forming a CDR hit library; selecting one of the FRs in the V_H or V_L region of the lead antibody; providing a second amino acid sequence that comprises at least 3 consecutive amino acid residues in the selected FR, the selected amino acid sequence being a FR lead sequence; comparing the FR lead sequence with a plurality of FR tester protein sequences; and selecting from the plurality of FR tester protein sequences at least two peptide segments that have at least 15% sequence identity with the FR lead sequence, the selected peptide segments forming a FR hit library; and combining the CDR hit library and the FR hit library to form a hit library.

According to the method, the plurality of CDR tester protein sequences may comprise amino acid sequences of human or non-human antibodies.

Also according to the method, the plurality of FR tester protein sequences may comprise amino acid sequences of human origins, preferably human or humanized antibodies (e.g., antibodies with at least 50% human sequence, preferably at least 70% human sequence, more preferably at least 90 % human sequence, and most preferably at least 95% human sequence in V_H or V_L), more preferably fully human antibodies, and most preferably human germline antibodies.

Also according to the method, at least one of the plurality of CDR tester protein sequences is different from the plurality of FR tester protein sequences.

Also according to the method, the plurality of CDR tester protein sequences are human or non-human antibody sequences and the plurality of FR tester protein sequences are human antibody sequences, preferably human germline antibody sequences.

The method may further comprise the step of: constructing a nucleic acid library comprising DNA segments encoding the amino acid sequences of the hit library.

Optionally, the method may further comprise the steps of: building an amino acid positional variant profile of the CDR hit library; converting the amino acid positional variant profile of the CDR hit library into a first nucleic acid positional variant profile by back-translating the amino acid positional variants into their corresponding genetic codons; and constructing a degenerate CDR nucleic acid library of DNA segments by combinatorially combining the nucleic acid positional variants.

Optionally, the genetic codons may be the ones that are preferred for expression in bacteria. Optionally, genetic codons may be the ones that can reduce the size chosen such that the diversity of the degenerate nucleic acid library of DNA segments is within the experimentally coverable diversity without undue experimental efforts, such as diversity below $1x10^7$, preferably below $1x10^6$.

In yet another embodiment, the method comprises the steps of: providing an amino acid sequence of the variable region of the heavy chain (V_H) or light chain (V_L) of a lead antibody; identifying the amino acid sequences in the FRs of the lead antibody; selecting one of the FRs in the V_H or V_L region of the lead antibody; providing a first amino acid sequence that comprises at least 3 consecutive amino acid residues in the selected FR, the selected amino acid sequence being a first FR lead sequence; comparing the first lead FR sequence with a plurality of FR tester protein sequences; and selecting from the plurality of FR tester protein sequences at least two peptide segments that have at least 15% sequence identity with the first FR lead sequence, the selected peptide segments forming a first FR hit library.

The method may further comprise the steps of: providing a second amino acid sequence that comprises at least 3 consecutive amino acid residues in a FR that is different from the selected FR, the selected amino acid sequence being a second FR lead sequence; comparing the second FR lead sequence with the plurality of FR tester protein sequences; and selecting from the plurality of FR tester protein sequences at least two

peptide segments that have at least 15% sequence identity with the second FR lead sequence, the selected peptide segments forming a second FR hit library; and combining the first FR hit library and the second FR hit library to form a hit library.

According to the method, the lead CDR sequence may comprise at least 5 consecutive amino acid residues in the selected CDR. The selected CDR may be selected from the group consisting of V_H CDR1, V_H CDR2, V_H CDR3, V_L CDR1, V_L CDR2, and V_L CDR3 of the lead antibody.

Also according to the method, the lead FR sequence may comprise at least 5 consecutive amino acid residues in the selected FR. The selected FR may be selected from the group consisting of V_H FR1, V_H FR2, V_H FR3, V_H FR4, V_L FR1, V_L FR2, V_L FR3 and V_L FR4 of the lead antibody.

The method may further comprise the step of: constructing a nucleic acid or degenerate nucleic acid library comprising DNA segments encoding the amino acid sequences of the hit library.

In another aspect of the invention, a method is provided for in silico selection of antibody sequences based on the amino acid sequence of a region in a lead antibody, i.e., the "lead sequence", and its 3D structure. The structure of the lead sequence is employed to search databases of protein structures for segments having similar 3D structures. These segments are aligned to yield a sequence profile, herein after referred to as the "lead sequence profile". The lead sequence profile is employed to search databases of protein sequences for remote homologues of the lead sequence having low sequence identity and yet structurally similar. By using the method, a library of diverse antibody sequences can be constructed and screened experimentally in vitro or in vivo for antibody mutants with improved or desired function(s).

In one embodiment, the method comprises the steps of: providing an amino acid sequence of the variable region of the heavy chain (V_H) or light chain (V_L) of a lead antibody; identifying the amino acid sequences in the CDRs of the lead antibody; selecting one of the CDRs in the V_H or V_L region of the lead antibody; providing an amino acid sequence that comprises at least 3 consecutive amino acid residues in the selected CDR, the selected amino acid sequence being a lead sequence; providing a three-dimensional structure of the lead sequence; building a lead sequence profile based on the

structure of the lead sequence; comparing the lead sequence profile with a plurality of tester protein sequences; and selecting from the plurality of tester protein sequences at least two peptide segments that have at least 10% sequence identity with lead sequence, the selected peptide segments forming a hit library.

According to the method, the three-dimensional structure of the lead sequence may be a structure derived from X-crystallography, nuclear magnetic resonance (NMR) spectroscopy or theoretical structural modeling.

According to the method, the step of building a lead sequence profile may include: comparing the structure of the lead sequence with the structures of a plurality of tester protein segments; determining the root mean square difference of the main chain conformations of the lead sequence and the tester protein segments; selecting the tester protein segments with root mean square difference of the main chain conformations less than 5 Å, preferably less than 4 Å, more preferably less than 3 Å, and most preferably less than 2 Å; and aligning the amino acid sequences of the selected tester protein segments with the lead sequence to build the lead sequence profile.

Optionally, the structures of the plurality of tester protein segments are retrieved from the protein data bank.

Optionally, the step of building a lead sequence profile may include: comparing the structure of the lead sequence with the structures of a plurality of tester protein segments; determining the Z-score of the main chain conformations of the lead sequence and the tester protein segments; selecting the segments of the tester protein segments with the Z-score higher than 2, preferably higher than 3, more preferably higher than 4, and most preferably higher than 5; and aligning the amino acid sequences of the selected tester protein segments with the lead sequence to build the lead sequence profile.

Optionally, the step of building a lead sequence profile may be implemented by an algorithm selected from the group consisting of CE, MAPS, Monte Carlo and 3D clustering algorithms.

The method may further comprise the step of: constructing a nucleic acid library comprising DNA segments encoding the amino acid sequences of the hit library.

Optionally, the method may further comprise the steps of: building an amino acid positional variant profile of the hit library; converting amino acid positional variant

profile of the hit library into a nucleic acid positional variant profile by back-translating the amino acid positional variants into their corresponding trinucleotide codons; and constructing a degenerate nucleic acid library of DNA segments by combinatorially combining the nucleic acid positional variants.

In yet another aspect of the invention, a method is provided for in silico selection of antibody sequences based on a 3D structure of a lead antibody. A lead sequence or sequence profile from a specific region of the lead antibody to be employed to search databases of protein sequences for remote homologues of the lead sequence having low sequence identity and yet structurally similar. These remote homologues form a hit library. The sequences in the hit library are subjected to evaluation for their structural compatibility with a 3D structure of the lead antibody, hereinafter referred to as the "lead structural template". Sequences in the hit library that are structurally compatible with the lead structural template are selected and screened experimentally in vitro or in vivo for antibody mutants with improved or desired function(s).

In one embodiment, the method comprises the steps of: providing an amino acid sequence of the variable region of the heavy chain (V_H) or light chain (V_L) of a lead antibody, the lead antibody having a known three dimensional structure which is defined as a lead structural template; identifying the amino acid sequences in the CDRs of the lead antibody; selecting one of the CDRs in the V_H or V_L region of the lead antibody; providing an amino acid sequence that comprises at least 3 consecutive amino acid residues in the selected CDR, the selected amino acid sequence being a lead sequence; comparing the lead sequence profile with a plurality of tester protein sequences; selecting from the plurality of tester protein sequences at least two peptide segments that have at least 10% sequence identity with lead sequence, the selected peptide segments forming a hit library; determining if a member of the hit library is structurally compatible with the lead structural template using a scoring function; and selecting the members of the hit library that score equal to or better thanor equal to the lead sequence.

According to the method, the scoring function is an energy scoring function selected from the group consisting of electrostatic interactions, van der Waals interactions, electrostatic solvation energy, solvent-accessible surface solvation energy, and conformational entropy.

Optionally, the scoring function is one incorporating a forcefield selected from the group consisting of the Amber forcefield, Charmm forcefield, the Discover cvff forcefields, the ECEPP forcefields, the GROMOS forcefields, the OPLS forcefields, the MMFF94 forcefield, the Tripos forcefield, the MM3 forcefield, the Dreiding forcefield,and UNRES forcefield,and other knowledge-based statistical forcefield (mean field) and structure-based thermodynamic potential functions.

Also according to the method, the step of selecting the members of the hit library includes selecting the members of the hit library that have a lower or equal total energy than that of the lead sequence calculated based on a formula of

$$\Delta E_{total} = E_{vdw} + E_{bond} + E_{angel} + E_{electrostatics} + E_{solvation}$$

Also according to the method, the step of selecting the members of the hit library includes selecting the members of the hit library that have a lower binding free energy than that of the lead sequence calculated as the difference between the bound and unbound states using a refined scoring function

$$\Delta G_b = \Delta G_{MM} + \Delta G_{sol} - T\Delta S_{ss}$$

where

$$\Delta G_{MM} = \Delta G_{ele} + \Delta G_{vdw} \quad (1)$$

$$\Delta G_{sol} = \Delta G_{ele\text{-}sol} + \Delta G_{ASA} \quad (2)$$

The method may further comprise the step of: constructing a nucleic acid library comprising DNA segments encoding the amino acid sequences of the hit library.

Optionally, the method may further comprise the steps of: building an amino acid positional variant profile of the hit library; converting amino acid positional variant profile of the hit library into a nucleic acid positional variant profile by back-translating the amino acid positional variants into their corresponding trinucleotide codons; and constructing a degenerate nucleic acid library of DNA segments by combinatorially combining the nucleic acid positional variants.

In yet another aspect of the invention, a method is provided for in silico selection of antibody sequences based on a 3D structure or structure ensemble of a lead antibody, or a structure ensemble of multiple antibodies, hereinafter collectively referred to as the lead structural template. A lead sequence or sequence profile from a specific region of the lead antibody to be employed to search databases of protein sequences for remote homologues of the lead sequence having low sequence identity and yet structurally similar. These remote homologues form a hit library. An amino acid positional variant profile (AA-PVP) of the hit library is built based on frequency of amino acid variant appearing at each position of the lead sequence. Based on the AA-PVP, a hit variant library is constructed by combinatorially combining the amino acid variant at each position of the lead sequence with or without cutoff of low frequency variants. The sequences in the hit variant library are subjected to evaluation for their structural compatibility with the lead structural template. Sequences in the hit library that are structurally compatible with the lead structural template are selected and screened experimentally in vitro or in vivo for antibody mutants with improved or desired function(s).

In one embodiment, the method comprises the steps of: providing an amino acid sequence of the variable region of the heavy chain (V_H) or light chain (V_L) of a lead antibody, the lead antibody having a known three dimensional structure which is defined as a lead structural template; identifying the amino acid sequences in the CDRs of the lead antibody; selecting one of the CDRs in the V_H or V_L region of the lead antibody; providing an amino acid sequence that comprises at least 3 consecutive amino acid residues in the selected CDR, the selected amino acid sequence being a lead sequence; comparing the lead sequence with a plurality of tester protein sequences; selecting from the plurality of tester protein sequences at least two peptide segments that have at least 10% sequence identity with lead sequence, the selected peptide segments forming a hit library; building an amino acid positional variant profile of the hit library based on frequency of amino acid variant appearing at each position of the lead sequence; combining the amino acid variants in the hit library to produce a combination of hit variants which form a hit variant library; determining if a member of the hit variant library is structurally compatible with the lead structural template using a scoring

function; and selecting the members of the hit variant library that score equal to or better than the lead sequence.

According to the method, the step of combining the amino acid variants in the hit library includes: selecting the amino acid variants with frequency of appearance higher than 2%, preferably 5%, more preferably 8% times, and most preferably 10% of the amino acid occurrence frequency for the cutoff and then include some of the amino acids from the lead sequence if they are missed after cutoff; and combining the selected amino acid variants in the hit library to produce a combination of hit variants which form a hit variant library.

According to the method, the scoring function is an energy scoring function selected from the group consisting of electrostatic interactions, van der Waals interactions, electrostatic solvation energy, solvent-accessible surface solvation energy, and conformational entropy.

Optionally, the scoring function is one incorporating a forcefield selected from the group consisting of the Amber forcefield, Charmm forcefield, the Discover cvff forcefields, the ECEPP forcefields, the GROMOS forcefields, the OPLS forcefields, the MMFF94 forcefield, the Tripos forcefield, the MM3 forcefield, the Dreiding forcefield,and UNRES forcefield,and other knowledge-based statistical forcefield (mean field) and structure-based thermodynamic potential functions.

The method may further comprise the step of: constructing a nucleic acid library comprising DNA segments encoding the amino acid sequences of the selected members of the hit variant library.

Optionally, the method may further comprise the steps of: partitioning theparsing the selected members of hit variant library into at least two sub-hit variant libraries; selecting a sub-hit variant library; building an amino acid positional variant profile of the selected sub-hit variant library; converting the amino acid positional variant profile of the selected sub-hit variant library into a nucleic acid positional variant profile by back-translating the amino acid positional variants into their corresponding trinucleotide codons; and constructing a degenerate nucleic acid library of DNA segments by combinatorially combining the nucleic acid positional variants.

The step of parsing the hit variant library may include: randomly selecting 10-30 members of the hit variant library that score equal to or better than the lead sequence, the selected members forming a sub-variant library.

Optionally, the step of parsing the hit variant library may include: building an amino acid positional variant profile of the hit variant library, resulting a hit variant profile; parsing the hit variant profile into segments of sub-variant profile based on the contact maps of the Cα, or Cβ or heavy atoms of the structure or structure ensembles of a lead sequence within certain distance cutoff (8A to 4.5 A). A structural model or lead structural template within a distance of 8 Å, preferably within 6 Å, more preferably within 5 Å, and most preferably within 4.5 Å.

In another embodiment, the method comprises the steps of: providing an amino acid sequence of the variable region of the heavy chain (V_H) or light chain (V_L) of a lead antibody, the lead antibody having a known three dimensional structure; providing 3D structures of one or more antibodies with different sequences in V_H or V_L region than that of the lead antibody; forming a structure ensemble by combining the structures of the lead antibody and the one or more antibodies; the structure ensemble being defined as a lead structural template; identifying the amino acid sequences in the CDRs of the lead antibody; selecting one of the CDRs in the V_H or V_L region of the lead antibody; providing an amino acid sequence that comprises at least 3 consecutive amino acid residues in the selected CDR, the selected amino acid sequence being a lead sequence; comparing the lead sequence with a plurality of tester protein sequences; selecting from the plurality of tester protein sequences at least two peptide segments that have at least 10% sequence identity with lead sequence, the selected peptide segments forming a hit library; building an amino acid positional variant profile of the hit library based on frequency of amino acid variant appearing at each position of the lead sequence; combining the amino acid variants in the hit library to produce a combination of hit variants which form a hit variant library; determining if a member of the hit variant library is structurally compatible with the lead structural template using a scoring function; and selecting the members of the hit variant library that score equal to or better than the lead sequence.

In a particular embodiment, the method comprises the steps of: a) providing an amino acid sequence of the variable region of the heavy chain (V_H) or light chain (V_L) of a lead antibody, the lead antibody having a known three dimensional structure; b) identifying the amino acid sequences in the CDRs of the lead antibody; c) selecting one of the CDRs in the V_H or V_L region of the lead antibody; d) providing an amino acid sequence that comprises at least 3 consecutive amino acid residues in the selected CDR, the selected amino acid sequence being defined as a lead sequence; e) comparing the lead sequence with a plurality of tester protein sequences; f) selecting from the plurality of tester protein sequences at least two peptide segments that have at least 10% sequence identity with lead sequence, the selected peptide segments forming a hit library; g) building an amino acid positional variant profile of the hit library based on frequency of amino acid variant appearing at each position of the lead sequence; h) combining the amino acid variants in the hit library to produce a combination of hit variants which form a hit variant library; I) determining if a member of the hit variant library is structurally compatible with the lead structural template using a scoring function; j) selecting the members of the hit variant library that score equal to or better than the lead sequence; k) constructing a degenerate nucleic acid library comprising DNA segments encoding the amino acid sequences of the selected members of the hit variant library; l) determining the diversity of the nucleic acid library, if the diversity is higher than $1x10^6$, repeating steps j) through l) until the diversity of the diversity of the nucleic acid library is equal to or lower than $1x10^6$; m) introducing the DNA segments in the degenerate nucleic acid library into cells of a host organism; n) expressing the DNA segments in the host cells such that recombinant antibodies containing the amino acid sequences of the hit library are produced in the cells of the host organism; o) selecting the recombinant antibody that binds to a target antigen with affinity higher than 10^6 M^{-1}; and p) repeating steps e) through o) if no recombinant antibody is found to bind to the target antigen with affinity higher than 10^6 M^{-1}.

In another particular embodiment, the method comprises the steps of:
a) providing an amino acid sequence of the variable region of the heavy chain (V_H) or light chain (V_L) of a lead antibody, the lead antibody having a known three dimensional structure which is defined as a lead structural template; b) identifying the amino acid

sequences in the CDRs of the lead antibody; c) selecting one of the CDRs in the V_H or V_L region of the lead antibody; d) providing an amino acid sequence that comprises at least 3 consecutive amino acid residues in the selected CDR, the selected amino acid sequence being defined as a lead sequence; e) mutating the lead sequence by substituting one or more of the amino acid residues of the lead sequence with one or more different amino acid residues, resulting in a lead sequence mutant library; f) determining if a member of the lead sequence mutant library is structurally compatible with the lead structural template using a first scoring function; g) selecting the lead sequence mutants that score equal to or better than the lead sequence; h) comparing the lead sequence with a plurality of tester protein sequences; I) selecting from the plurality of tester protein sequences at least two peptide segments that have at least 10% sequence identity with lead sequence, the selected peptide segments forming a hit library; j) building an amino acid positional variant profile of the hit library based on frequency of amino acid variant appearing at each position of the lead sequence; k) combining the amino acid variants in the hit library to produce a combination of hit variants; l) combining the selected lead sequence mutants with the combination of hit variants to produce a hit variant library; m) determining if a member of the hit variant library is structurally compatible with the lead structural template using a second scoring function; n) selecting the members of the hit variant library that score equal to or better than the lead sequence; o) constructing a degenerate nucleic acid library comprising DNA segments encoding the amino acid sequences of the selected members of the hit variant library; p) determining the diversity of the nucleic acid library, and if the diversity is higher than $1x10^6$, repeating steps n) through p) until the diversity of the diversity of the nucleic acid library is equal to or lower than $1x10^6$; q) introducing the DNA segments in the degenerate nucleic acid library into cells of a host organism; r) expressing the DNA segments in the host cells such that recombinant antibodies containing the amino acid sequences of the hit library are produced in the cells of the host organism; s) selecting the recombinant antibody that binds to a target antigen with affinity higher than 10^6 M^{-1}; and t) repeating steps e) through s) if no recombinant antibody is found to bind to the target antigen with affinity higher than 10^6 M^{-1}.

In yet another aspect of the present invention, a computer-implemented method is provided for constructing a library of mutant antibodies based on a lead antibody.

In one embodiment, the method comprises: taking as an input an amino acid sequence that comprises at least 3 consecutive amino acid residues in a CDR region of the lead antibody, the amino acid sequence being a lead sequence; employing a computer executable logic to compare the lead sequence with a plurality of tester protein sequences; selecting from the plurality of tester protein sequences at least two peptide segments that have at least 15% sequence identity with lead sequence; and generating as an output the selected peptide segments which form a hit library.

According to any of the above methods, the length of the lead sequence is preferably between 5-100 aa, more preferably between 6-80 aa, and most preferably between 8-50 aa.

According to any of the above methods, the step of identifying the amino sequences in the CDRs is carried out by using Kabat criteria or Chothia criteria.

Also according to any of the above methods, the lead sequence may comprise an amino acid sequence from a particular region within the V_H or V_L of the lead antibody, CDR1, CDR2 or CDR3, or from a combination of the CDR and FRs, such as CDR1-FR2, FR2-CDR2-FR3, and the full-length V_H or V_L sequence. The lead sequence preferably comprises at least 6 consecutive amino acid residues in the selected CDR, more preferably at least 7 consecutive amino acid residues in the selected CDR, and most preferably all of the amino acid residues in the selected CDR.

Also according to any of the above methods, the lead sequence may further comprise at least one of the amino acid residues immediately adjacent to the selected CDR.

Also according to any of the above methods, the lead sequence may further comprise at least one of the FRs flanking the selected CDR.

Also according to any of the above methods, the lead sequence may further comprise one or more CDRs or FRs adjacent the C-terminus or N-terminus of the selected CDR.

Also according to any of the above methods, the lead structural template may be a 3D structure of a fully assembled lead antibody, or a heavy chain or light chain variable region of the lead antibody (e.g., CDR, FR and a combination thereof).

Also according to any of the above methods, the plurality of tester protein sequences includes preferably antibody sequences, more preferably human antibody sequences, and most preferably human germline antibody sequences (V-database), especially for the framework regions.

Also according to any of the above methods, the plurality of tester protein sequences is retrieved from genbank of the NIH or Swiss-Prot database or the Kabat database for CDRs of antibodies.

Also according to any of the above methods, the step of comparing the lead sequence with the plurality of tester protein sequences is implemented by an algorithm selected from the group consisting of BLAST, PSI-BLAST, profile HMM, and COBLATH.

Also according to any of the above methods, the sequence identity of the selected peptide segments in the hit library with the lead sequence is preferably at least 25%, preferably at least 35%, and most preferably at least 45%.

According to any of the above method, the method further comprises the following steps: introducing the DNA segments in the nucleic acid or degenerate nucleic acid library into cells of a host organism; expressing the DNA segments in the host cells such that recombinant antibodies containing the amino acid sequences of the hit library encoded by the nucleic acid or degenerate nucleic acid library are produced in the cells of the host organism; and selecting the recombinant antibody that binds to a target antigen with affinity higher than 10^6 M^{-1}, optionally 10^7 M^{-1}, optionally 10^8 M^{-1}, optionally 10^9 M^{-1}, optionally $2x10^9$ M^{-1}, optionally $5x10^9$ M^{-1}, optionally $1x10^{10}$ M^{-1}, optionally $5x10^{10}$ M^{-1}, and optionally $1x10^{11}$ M^{-1}.

The recombinant antibodies may be fully assembled antibodies, Fab fragments, Fv fragments, or single chain antibodies.

The host organism includes any organism or its cell line that is capable of expressing transferred foreign genetic sequence, including but not limited to bacteria, yeast, plant, insect, and mammals.

The recombinant antibodies may be fully assembled antibodies, Fab fragments, Fv fragments, or single chain antibodies. For example, the recombinant antibodies may be expressed in bacterial cells and displayed on the surface of phage particles. The

recombinant antibodies displayed on phage particles may be a double-chain heterodimer formed between V_H and V_L. The heterodimerization of V_H and V_L chains may be facilitated by a heterodimer formed between two non-antibody polypeptide chains fused to the V_H and V_L chains, respectively. For example, these two non-antibody polypeptide may be derived from a heterodimeric receptors $GABA_B$ R1 (GR1) and R2 (GR2), respectively.

Alternatively, the recombinant antibodies displayed on phage particles may be a single-chain antibody containing V_H and V_L linked by a peptide linker. The display of the single chain antibody on the surface of phage particles may be facilitated by a heterodimer formed between a fusion of the single chain antibody with GR1 and a fusion of phage pIII capsid protein with GR2.

The target antigen to be screened against includes small molecules and macromolecules such as proteins, peptides, nucleic acids and polycarbohydrates.

Any of the above methods may further comprise the following steps: introducing the DNA segments in the nucleic acid or degenerate nucleic acid library into cells of a host organism; expressing the DNA segments in the host cells such that recombinant antibodies containing the amino acid sequences of the hit library are produced in the cells of the host organism; and selecting the recombinant antibody that binds to a target antigen with affinity higher than 10^6 M^{-1}.

The binding affinity of the selected recombinant antibody to the target antigen is optionally higher than 10^7 M^{-1}, optionally higher than 10^8 M^{-1}, optionally higher than 1 x 10^9 M^{-1}, optionally higher than 2 x 10^9 M^{-1}, optionally higher than 5 x 10^9 M^{-1}, optionally higher than 8 x 10^9 M^{-1}, optionally higher than 1 x 10^{10} M^{-1}, optionally higher than 2 x 10^{10} M^{-1}, optionally higher than 5 x 10^{10} M^{-1}, optionally higher than 8 x 10^{10} M^{-1}, or optionally higher than 1 x 10^{11} M^{-1}.

The binding affinity of the selected antibody to its antigen may vary, depending the form of antibody being tested. The selected antibody being tested may be in the form of a single-chain antibody (scFv) comprising V_H and V_L designed by using the methodology of the present invention. Optionally, the selected antibody being tested may be in the form of a Fab comprising V_H and V_L designed by using the methodology of the present invention. Presumably due to its higher conformational flexibility and instability,

the binding affinity of the selected antibody in the form of scFv may be 1-2 magnitude lower than that in the form of Fab. Accordingly, for a selected antibody in the form of scFv, the dissociation constant is preferably lower than 10^6 M^{-1}, optionally higher than 10^7 M^{-1}, optionally higher than 1 x 10^8 M^{-1}, optionally higher than 2 x 10^8 M^{-1}, optionally higher than 5 x 10^8 M^{-1}, optionally higher than 8 x 10^8 M^{-1}, optionally higher than 1 x 10^9 M^{-1}, optionally higher than 2 x 10^9 M^{-1}, optionally higher than 5 x 10^9 M^{-1}, optionally higher than 8 x 10^9 M^{-1}, optionally higher than 1 x 10^{10} M^{-1}, optionally higher than 5 x 10^{10} M^{-1}, or optionally higher than 1x 10^{11} M^{-1}.

The binding affinity of the selected recombinant antibody to the target antigen may also be represented by the dissociation constant K_d measured in a kinetic study of the binding interaction between the antibody and the target antigen at a certain temperature (e.g., 4°C, 25°C, 35°C, 37°C, or 42°C), for example, by using an instrument such as a BIAcore biosensor (*see* EXAMPLE). Generally, the lower K_d measured, the higher affinity the antibody; and the higher temperature of measurement, the higher K_d measured for the same antibody.

The K_d of a selected antibody is optionally lower than 100 nM, optionally lower than 10 nM, optionally lower than 8 nM, optionally lower than 8 nM, optionally lower than 5 nM, optionally lower than 1 nM, optionally lower than 0.8 nM, optionally lower than 0.5 nM, optionally lower than 0.2 nM, optionally lower than 0.1 nM, optionally lower than 0.08 nM, optionally lower than 0.05 nM, optionally lower than 0.01 nM, or optionally lower than 0.005 nM, in the form of scFv, Fab, or other form of antibody measured at a temperature of about 4°C, 25°C, 35°C, 37°C or 42°C.

According to any of the above embodiments, the designed proteins (e.g.. antibodies) may be synthesized, or expressed in cells of any organism, including but not limited to bacteria, yeast, plant, insect, and mammal. Particular types of cells include, but are not limited to, *Drosophila melanogaster* cells, *Saccharomyces cerevisiae* and other yeasts, E. coli, Bacillus subtilis, SF9 cells, C129 cells, 293 cells, Neurospora, BHK, CHO, COS, and HeLa cells, fibroblasts, Schwanoma cell lines, immortalized mammalian myeloid and lymphoid cell lines, Jurkat cells, mast cells and other endocrine and exocrine cells, and neuronal cells. Examples of mammalian cells include, but are not limited to, tumor cells of all types (particularly melanoma, myeloid leukemia, carcinomas of the

lung, breast, ovaries, colon, kidney, prostate, pancreas and testes), cardiomyocytes, endothelial cells, epithelial cells, lymphocytes (T-cell and B cell), mast cells, eosinophils, vascular intimal cells, hepatocytes, leukocytes including mononuclear leukocytes, stem cells such as haemopoetic, neural, skin, lung, kidney, liver and myocyte stem cells, osteoclasts, chondrocytes and other connective tissue cells, keratinocytes, melanocytes, liver cells, kidney cells, and adipocytes.

Preferably, the designed protein is purified or isolated after expression according to methods known to those skilled in the art. Examples of purification methods include electrophoretic, molecular, immunological and chromatographic techniques, including ion exchange, hydrophobic, affinity, and reverse-phase HPLC chromatography, and chromatofocusing. The degree of purification necessary will vary depending on the use of the designed protein. In some instances no purification will be necessary.

Also according to any of the embodiments described above, the designed proteins can be screened for a desired function, preferably a biological function such as their binding to a known binding partner, physiological activity, stability profile (pH, thermal, buffer conditions), substrate specificity, immunogenicity, toxicity, etc.

In the screening using a cell-based assay, the designed protein may be selected based on an altered phenotype of the cell, preferably in some detectable and/or measurable way. Examples of phenotypic changes include, but are not limited to, gross physical changes such as changes in cell morphology, cell growth, cell viability, adhesion to substrates or other cells, and cellular density; changes in the expression of one or more RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes in the equilibrium state (i.e. half-life) or one or more RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes in the localization of one or more RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes in the bioactivity or specific activity of one or more RNAs, proteins, lipids, hormones, cytokines, receptors, or other molecules; changes in the secretion of ions, cytokines, hormones, growth factors, or other molecules; alterations in cellular membrane potentials, polarization, integrity or transport; changes in infectivity, susceptability, latency, adhesion, and uptake of viruses and bacterial pathogens.

According to any of the above embodiment, the designed proteins (e.g.. antibodies) may be synthesized, or expressed as fusion proteins with a tag protein or peptide. The tag protein or peptide may be used to identify, isolate, signal, stabilize, increase flexibility of, increase degradation of, increase secretion, translocation or intracellular retention or enhance expression of the designed proteins.

The invention further provides: isolated nucleic acid encoding the antibody; a vector comprising that nucleic acid, optionally operably linked to control sequences recognized by a host cell transformed with the vector; a host cell comprising that vector; a process for producing the antibody comprising culturing the host cell so that the nucleic acid is expressed and, optionally, recovering the antibody from the host cell culture.

3. Method of Using the Antibodies of the Present Invention

The antibodies designed by using the inventive methods may be used for diagnosing or therapeutic treatment of various diseases, including but not limited to, cancer, autoimmune diseases such as multiple sclerosis, rheumatoid arthritis, systemic lupus erythematosus, Type I diabetes, and myasthenia gravis, graft-versus-host disease, cardiovascular diseases, viral infection such as HIV, hepatitis viruses, and herpes simplex virus, bacterial infection, allergy, Type II diabetes, hematological disorders such as anemia.

The antibodies can also be used as conjugates that are linked with diagnostic or therapeutic moieties, or in combination with chemotherapeutic or biological agents.

The antibodies of the invention may be used as affinity purification agents. In this process, the antibodies are immobilized on a solid phase such a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody is contacted with a sample containing the VEGF protein (or fragment thereof) to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the VEGF protein, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent, such as glycine buffer, pH 5.0, that will release the VEGF protein from the antibody.

Anti-VEGF antibodies may also be useful in diagnostic assays for VEGF protein, e.g., detecting its expression in specific cells, tissues, or serum. Such diagnostic methods may be useful in cancer diagnosis.

For diagnostic applications, the antibody may be labeled with a detectable moiety. For example, the antibody can be labeled with a radioisotope, such as ^{35}S, ^{14}C, ^{125}I, ^{3}H, and ^{131}I, using the techniques described in Current Protocols in Immunology, Volumes 1 and 2, Coligen et al., Ed. Wiley-Interscience, New York, New York, Pubs. (1991) for example and radioactivity can be measured using scintillation counting.

The antibody may also be conjugated with a fluorescent label such as rare earth chelates (e.g., europium chelates), fluorescein and its derivatives, rhodamine and its derivatives, dansyl, Lissamine, phycoerythrin and Texas Red. The fluorescent labels can be conjugated to the antibody using the techniques disclosed in Current Protocols in Immunology, *supra*, for example. Fluorescence can be quantified using a fluorimeter.

The antibody can be labeled with various enzyme-substrate labels such as those disclosed in U.S. Pat. No. 4,275,149. The enzyme generally catalyzes a chemical alteration of the chromogenic substrate which can be measured using various techniques. For example, the enzyme may catalyze a color change in a substrate, which can be measured spectrophotometrically. Alternatively, the enzyme may alter the fluorescence or chemiluminescence of the substrate.

In one diagnostic application, the invention provides a method for determining the presence of VEGF protein comprising exposing a sample suspected of containing the VEGF protein to the anti-VEGF antibody and determining binding of the antibody to the sample. For this use, the invention provides a kit comprising the antibody and instructions for using the antibody to detect the VEGF protein.

The antibodies of the present invention can also be formulated for delivery via a wide variety of routes of administration. Therapeutic formulations of the antibody are prepared for storage by mixing the antibody having the desired degree of purity with optional physiologically acceptable carriers, excipients or stabilizers (Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)), in the form of lyophilized formulations or aqueous solutions. For example, the antibodies may be administered or coadministered orally, topically, parenterally, intraperitoneally, intravenously,

intraarterially, transdermally, sublingually, intramuscularly, rectally, transbuccally, intranasally, via inhalation, vaginally, intraoccularly, via local delivery (for example by a catheter or a stent), subcutaneously, intraadiposally, intraarticularly, or intrathecally.

Depending on the type and severity of the disease, about 1 ug/kg to about 50 mg/kg (e.g., 0.1-20 mg/kg, 0.5-15 mg/Kg, and 1-10 mg/kg) of antibody is an initial candidate dosage for administration to the patient, whether, for example, by one or more separate administrations, or by continuous infusion. A typical daily or weekly dosage might range from about 1 ug/kg to about 20 mg/kg or more, depending on the factors mentioned above. For repeated administrations over several days or longer, depending on the condition, the treatment is repeated until a desired suppression of disease symptoms occurs. However, other dosage regimens may be useful. The progress of this therapy is easily monitored by conventional techniques and assays, including, for example, radiographic tumor imaging.

The anti-VEGF antibodies of the present invention may be used to treat a wide variety of indications for anti-VEGF antibodies have therapeutic activity. Such indications include, but are not limited to, restenosis (e.g. coronary, carotid, and cerebral lesions), benign tumors, a various types of cancers such as primary tumors and tumor metastasis, abnormal stimulation of endothelial cells (atherosclerosis), insults to body tissue due to surgery, abnormal wound healing, abnormal angiogenesis, diseases that produce fibrosis of tissue, muscular degeneration, repetitive motion disorders, disorders of tissues that are not highly vascularized, and proliferative responses associated with organ transplants.

Examples of benign tumors include hemangiomas, hepatocellular adenoma, cavernous haemangioma, focal nodular hyperplasia, acoustic neuromas, neurofibroma, bile duct adenoma, bile duct cystanoma, fibroma, lipomas, leiomyomas, mesotheliomas, teratomas, myxomas, nodular regenerative hyperplasia, trachomas and pyogenic granulomas.

Specific types of cancers include, but are not limited to, leukemia, breast cancer, skin cancer, bone cancer, prostate cancer, liver cancer, lung cancer, brain cancer, cancer of the larynx, gallbladder, pancreas, rectum, parathyroid, thyroid, adrenal, neural tissue, head and neck, colon, stomach, bronchi, kidneys, basal cell carcinoma, squamous cell

carcinoma of both ulcerating and papillary type, metastatic skin carcinoma, osteo sarcoma, Ewing's sarcoma, veticulum cell sarcoma, myeloma, giant cell tumor, small-cell lung tumor, gallstones, islet cell tumor, primary brain tumor, acute and chronic lymphocytic and granulocytic tumors, hairy-cell tumor, adenoma, hyperplasia, medullary carcinoma, pheochromocytoma, mucosal neuronms, intestinal ganglloneuromas, hyperplastic corneal nerve tumor, marfanoid habitus tumor, Wilm's tumor, seminoma, ovarian tumor, leiomyomater tumor, cervical dysplasia and in situ carcinoma, neuroblastoma, retinoblastoma, soft tissue sarcoma, malignant carcinoid, topical skin lesion, mycosis fungoide, rhabdomyosarcoma, Kaposi's sarcoma, osteogenic and other sarcoma, malignant hypercalcemia, renal cell tumor, polycythermia vera, adenocarcinoma, glioblastoma multiforma, leukemias, lymphomas, malignant melanomas, epidermoid carcinomas, and other carcinomas and sarcomas.

Diseases associated with abnormal angiogenesis include, but are not limited to, rheumatoid arthritis, ischemic-reperfusion related brain edema and injury, cortical ischemia, ovarian hyperplasia and hypervascularity, (polycystic ovary syndrom), endometriosis, psoriasis, diabetic retinopaphy, and other ocular angiogenic diseases such as retinopathy of prematurity (retrolental fibroplastic), macular degeneration, corneal graft rejection, neuroscular glaucoma and Oster Webber syndrome. In a particular embodiment, the anti-VEGF antibodies of the present invention can be used for treating age-related macular degeneration (AMD).

Examples of retinal/choroidal neovascularization include, but are not limited to, Bests diseases, myopia, optic pits, Stargarts diseases, Pagets disease, vein occlusion, artery occlusion, sickle cell anemia, sarcoid, syphilis, pseudoxanthoma elasticum carotid abostructive diseases, chronic uveitis/vitritis, mycobacterial infections, Lyme's disese, systemic lupus erythematosis, retinopathy of prematurity, Eales disease, diabetic retinopathy, macular degeneration, Bechets diseases, infections causing a retinitis or chroiditis, presumed ocular histoplasmosis, pars planitis, chronic retinal detachment, hyperviscosity syndromes, toxoplasmosis, trauma and post-laser complications, diseases associated with rubesis (neovascularization of the ankle) and diseases caused by the abnormal proliferation of fibrovascular or fibrous tissue including all forms of proliferative vitreoretinopathy.

Examples of corneal neovascularization include, but are not limited to, epidemic keratoconjunctivitis, Vitamin A deficiency, contact lens overwear, atopic keratitis, superior limbic keratitis, pterygium keratitis sicca, sjogrens, acne rosacea, phylectenulosis, diabetic retinopathy, retinopathy of prematurity, corneal graft rejection, Mooren ulcer, Terrien's marginal degeneration, marginal keratolysis, polyarteritis, Wegener sarcoidosis, Scleritis, periphigoid radial keratotomy, neovascular glaucoma and retrolental fibroplasia, syphilis, Mycobacteria infections, lipid degeneration, chemical burns, bacterial ulcers, fungal ulcers, Herpes simplex infections, Herpes zoster infections, protozoan infections and Kaposi sarcoma.

The anti-VEGF antibodies of the present invention may be used in combination with an anti-angiogenesis agent for the treatment of diseases associated with abnormal angiogenesis.

Examples of anti-angiogenesis agents include, but are not limited to, retinoid acid and derivatives thereof, 2-methoxyestradiol, ANGIOSTATIN™ protein, ENDOSTATIN™ protein, suramin, squalamine, tissue inhibitor of metalloproteinase-I, tissue inhibitor of metalloproteinase-2, plasminogen activator inhibitor-1, plasminogen activator inhibitor-2, cartilage-derived inhibitor, paclitaxel, platelet factor 4, protamine sulphate (clupeine), sulphated chitin derivatives (prepared from queen crab shells), sulphated polysaccharide peptidoglycan complex (sp-pg), staurosporine, modulators of matrix metabolism, including for example, proline analogs ((l-azetidine-2-carboxylic acid (LACA), cishydroxyproline, d,l-3,4-dehydroproline, thiaproline], α, α -dipyridyl, β-aminopropionitrile fumarate, 4-propyl-5-(4-pyridinyl)-2(3h)-oxazolone; methotrexate, mitoxantrone, heparin, interferons, 2 macroglobulin-serum, chimp-3, chymostatin, beta.-cyclodextrin tetradecasulfate, eponemycin; fumagillin, gold sodium thiomalate, d-penicillamine (CDPT), beta.-1-anticollagenase-serum, □2-antiplasmin, bisantrene, lobenzarit disodium, n-(2-carboxyphenyl-4-chloroanthronilic acid disodium or "CCA", thalidomide; angostatic steroid, cargboxynaminolmidazole; metalloproteinase inhibitors such as BB94. Other anti-angiogenesis agents include antibodies, such as monoclonal antibodies against these angiogenic growth factors: bFGF, aFGF, FGF-5, VEGF isoforms, VEGF-C, HGF/SF and Ang-1/Ang-2.

The anti-VEGF antibodies of the present invention, preferably those having therapeutic synergistic effects with the anti-VEGF antibodies, may be employed in combination with the anti-VEGF antibodies to further enhance the therapeutic effects of these two types of drug. Examples of the therapeutic agent include, but are not limited to, alkylating agents, antibiotic agents, antimetabolic agents, hormonal agents, plant-derived agents, and biologic agents.

Examples of alkylating agents include, but are not limited to, bischloroethylamines (nitrogen mustards, e.g. chlorambucil, cyclophosphamide, ifosfamide, mechlorethamine, melphalan, uracil mustard), aziridines (e.g. thiotepa), alkyl alkone sulfonates (e.g. busulfan), nitrosoureas (e.g. carmustine, lomustine, streptozocin), nonclassic alkylating agents (altretamine, dacarbazine, and procarbazine), platinum compounds (carboplastin and cisplatin).

Examples of antibiotic agents include, but are not limited to, anthracyclines (e.g. doxorubicin, daunorubicin, epirubicin, idarubicin and anthracenedione), mitomycin C, bleomycin, dactinomycin, plicatomycin.

Examples of antimetabolic agents include, but are not limited to, fluorouracil (5-FU), floxuridine (5-FUdR), methotrexate, leucovorin, hydroxyurea, thioguanine (6-TG), mercaptopurine (6-MP), cytarabine, pentostatin, fludarabine phosphate, cladribine (2-CDA), asparaginase, imatinib mesylate (or GLEEVAC®), and gemcitabine.

Examples of such hormonal agents are synthetic estrogens (e.g. diethylstibestrol), antiestrogens (e.g. tamoxifen, toremifene, fluoxymesterol and raloxifene), antiandrogens (bicalutamide, nilutamide, flutamide), aromatase inhibitors (e.g., aminoglutethimide, anastrozole and tetrazole), ketoconazole, goserelin acetate, leuprolide, megestrol acetate and mifepristone.

Examples of plant-derived agents include, but are not limited to, vinca alkaloids (e.g., vincristine, vinblastine, vindesine, vinzolidine and vinorelbine), podophyllotoxins (e.g., etoposide (VP-16) and teniposide (VM-26)), camptothecin compounds (e.g., 20(S) camptothecin, topotecan, rubitecan, and irinotecan), taxanes (e.g., paclitaxel and docetaxel).

Examples of biologic agents include, but are not limited to, immuno-modulating proteins such as cytokines, monoclonal antibodies against tumor antigens, tumor

suppressor genes, and cancer vaccines. Examples of interleukins that may be used in conjunction with the antibody of the present invention include, but are not limited to, interleukin 2 (IL-2), and interleukin 4 (IL-4), interleukin 12 (IL-12). Examples of interferons that may be used in conjunction with the antibody of the present invention include, but are not limited to, interferon a, interferon b (fibroblast interferon) and interferon g (fibroblast interferon). Examples of such cytokines include, but are not limited to erythropoietin (epoietin a), granulocyte-CSF (filgrastin), and granulocyte, macrophage-CSF (sargramostim). Other immuno-modulating agents other than cytokines include, but are not limited to bacillus Calmette-Guerin, levamisole, and octreotide.

The anti-VEGF antibodies of the present invention may also be combined with a tumor necrosis factor (TNF) or its mutein in the treatment of the above diseases or conditions. The administration of anti-VEGF antibody and TNF is repeated until the desired clinical effect is achieved. In instances where solid tumors are found in the limbs or in other locations susceptible to isolation from the general circulation, the antibody and/or TNF may be administered to the isolated tumor or organ. In other embodiments, a FGF or platelet-derived growth factor (PDGF) antagonist, such as an anti-FGF or an anti-PDGF neutralizing antibody, is administered to the patient in conjunction with the anti-VEGF antibody. Treatment with anti-VEGF antibodies optimally may be suspended during periods of wound healing or desirable neovascularization.

Example of monoclonal antibodies against tumor antigens that can be used in conjunction with the anti-VEGF antibodies of the present invention include, but are not limited to, HERCEPTIN® (Trastruzumab), RITUXAN® (Rituximab), MYLOTARG® (gemtuzumab ozogamicin), CAMPATH® (alemtuzumab), ZEVALIN® (ibritumomab yiuxetan), PANOREX® (edrecolomab), BEXXAR® (tositumomab), ERBITUX® (cetuximab), and AVASTIN® (bevacizumab).

Examples of the tumor suppressor genes include, but are not limited to, *DPC-4, NF-1, NF-2, RB, p53, WT1, BRCA1* and *BRCA2*.

Example of cancer vaccines include, but are not limited to gangliosides (GM2), prostate specific antigen (PSA), □-fetoprotein (AFP), carcinoembryonic antigen (CEA) (produced by colon cancers and other adenocarcinomas, e.g. breast, lung, gastric, and pancreas cancer s), melanoma associated antigens (MART-1, gp100, MAGE 1,3

tyrosinase), papillomavirus E6 and E7 fragments, whole cells or portions/lysates of antologous tumor cells and allogeneic tumor cells.

An adjuvant may be used to augment the immune response to TAAs. Examples of adjuvants include, but are not limited to, bacillus Calmette-Guerin (BCG), endotoxin lipopolysaccharides, keyhole limpet hemocyanin (GKLH), interleukin-2 (IL-2), granulocyte-macrophage colony-stimulating factor (GM-CSF) and cytoxan, a chemotherapeutic agent which is believe to reduce tumor-induced suppression when given in low doses.

EXAMPLES

1. In Silico design of Anti-VEGF Antibodies for Humanization and Affinity Maturation

The methodology provided in the present invention was used to design libraries for humanization of non-human antibodies and to optimize the affinity and other attributes of antibodies. Novel variants of amino acids and nucleic acids of antibodies were generated with human or human like sequences while their binding affinity, stability, expression are improved significantly.

The inventive process was carried out computationally in a high throughput manner by mining the ever-expanding databases of protein sequences of all organisms, especially human and by relating their specific sequences or their variants with functional enhancement such as binding affinity and stability that are tested experimentally. By using the inventive methodology, an expanded and yet functionally biased library of antibodies were constructed based on computational evaluation of extremely diverse protein sequences and functionally relevant structures *in silico* and subsequently tested by experimental screening and selection in vitro or in vivo.

In general, the method was implemented in a computer through *in silico* selection of protein sequences based on the amino acid sequence of a target structural/functional motif or domain in a lead protein, herein after referred to as the "lead sequence". The lead sequence was employed to search databases of protein sequences. The choice of the database depends on the specific functional requirement of the designed motifs. For example, if the lead protein is an enzyme and the target motif includes the active site of the enzyme, databases of proteins/peptides of a particular origin, organism, species or combinations thereof, may be queried using various search criteria to yield a hit list of sequences each of which can substitute the target motif in the lead protein. A similar approach may be used for designing other motifs or domains of the lead protein. The designed sequences for each individual motif/domain may be combined to generate a library of designed proteins. In addition, to reduce immunogenicity of the designed proteins for human applications such as therapeutics or diagnosis, databases of proteins

of human origin or humanized proteins are preferably searched to yield the hit list of sequences, especially for motifs derived from sites of the lead protein that serves as the scaffolding of the lead proteins such as the frameworks of an antibody. The library of designed proteins can be tested experimentally to yield proteins with improved biological function(s) over the lead protein.

In this example, the inventive methodology was implemented in designing anti-VEGF antibodies that are diverse in sequence and yet functionally related to each other. Based on the designed antibody sequences, a library of antibodies were constructed to include diverse sequences in the complementary determining regions (CDRs) and/or humanized frameworks (FRs) of a non-human antibody in a high throughput manner. This library of antibodies were screened against human VEGF for improved function such as binding affinity and pharmacokinetic property.

In designing the anti-VEGF antibodies, in silico selection of antibody sequences was based on the amino acid sequence of a region or motif in a lead anti-VEGF antibody, herein after referred to as the "lead sequence". The lead sequence was employed to search databases of protein sequences. The choice of the database depends on the specific functional requirement of the designed motifs. For example: in order to design the framework regions of variable chains for therapeutic application, collections of protein sequences that are evolutionarily related such as fully human immunoglobulin sequences and human germline immunoglobulin sequences would be used except for a few structurally critical sites. This would reduce the immunogenic response by preserving the origin of the sequences by introducing as few foreign mutants as possible in this highly conserved region (for framework regions). On the other hand, diverse sequence databases such as immunoglobulin sequences of various species or even unrelated sequence in genbank can be used to design the CDRs in order to improve binding affinity with antigens in this highly variable region. By using the method, a library of diverse antibody sequences were constructed and screened experimentally in vitro and/or in vivo for antibody mutants with improved or desired function(s).

1) Anti-VEGF antibody libraries designed *in silico* for affinity maturation

The amino acid sequence of the variable heavy chain regions of a murine anti-VEGF antibody and the segments for frameworks and CDRs (underlined and annotated) is:

EIQLVQSGPELKQPGETVRISCKASGYTFTNYGMN(VH/CDR1)WVKQAPGKGLK
WMGWINTYTGEPTYAADFKR(VH/CDR2)RFTFSLETSASTAYLQISNLKNDDTAT
YFCAKYPHYYGSSHWYFDV(VH/CDR3)WGAGTTVTVSS (SEQ ID NO:283)

This V_H sequence therein after referred to as the "parental anti-VEGF antibody". The frameworks and CDRs are designated according to the Kabat criteria (Kabat EA, Redi-Miller M, Perry HM, Gottesman KS (1987) Sequences of Proteins of Immunological Interest 4th edit, National Institutes of Health, Bethesda, MD).

The CDR and framework regions of the antibody were targeted using a modular in silico evolutionary design approach as described in more detail in U.S. Patent Application Serial Nos: 10/443,134, 10/153,159, 10/153,176, 10/125,687, and 60/284,407, which are incorporated herein by reference in their entirety. Using murine anti-VEGF antibody as the lead protein and its V_H CDR3 as the lead sequence, digital libraries of V_H CDR3 were constructed by following the procedure.

The lead sequence includes V_H CDR3 of the parental anti-VEGF antibody and a few amino acid residues from the adjacent framework regions CAKYPHYYGSSHWYFDVWG. A hit library was constructed by searching and selecting hit amino acid sequences to V_H CDR3 from a sequence database. Variant profile was built to list all variants at each position based on the hit library and filtered with certain cutoff value to reduce of the size of the resulting hit variant library within computational or experimental limit. Variant profiles were also built in order to facilitate i) the sampling of the sequence space that covers the preferred region in the fitness landscape; ii) the partitioning and synthesis of degenerate nucleic acid libraries that target the preferred peptide ensemble sequences; iii) the experimental screening of the antibody libraries for the desired function; and iv) the analysis of experimental results with feedback for further design and optimization.

The lead structural templates were obtained from the available X-ray structures of the complexes formed between VEGF and anti-VEGF antibodies. The complex structure of VEGF and parental anti-VEGF antibody is designated as 1BJ1, and that formed

between VEGF and matured anti-VEGF antibody 1CZ8. The results from 1CZ8 structural template were similar to those from 1BJ1 in the relative ranking order of the scanned sequences. Structural models of anti-VEGF antibodies can be also used.

The lead sequence for V_H CDR3 is taken from the parental anti-VEGF antibody according to Kabat classification with amino acid residues CAK and WG from the adjacent framework regions flanking the V_H CDR3 sequence at N- and C- terminus, respectively. Only V_H CDR3 sequence of the parental antibody was used to build the HMM for searching the protein databases.

The HMM built using the single lead sequence or sequences of the structural ensembles was calibrated and used to search the Kabat database (Johnson, G and Wu, TT (2001) Nucleic Acids Research, 29, 205-206). All sequence hits that are above expectation value or E-value are listed and aligned using HAMMER 2.1.1 package. After removing the redundant sequences from the hit list, the remaining hit sequences for the lead HMM form the hit library.

The variant profile at each position was used to build the AA-PVP table (amino acid positional variant profile), which gives the number of occurrence of each amino acid residue at each position.

The variant profile can filtered to remove variants that occur at or less than the certain cutoff frequency and/or in combination with variant reprofiling using structure-based scoring. The variant profile from the sequence pool provides informative data to identify the positions in the lead sequence that can be either varied or fixed. The sites can be divided into three categories: i) Structurally conserved sites remain conserved over evolution. The high frequency residues can be used to maintain the scaffold of the target motif at these positions; ii) variable functional hot spots should be targeted with focused mutagenesis; iii) combination of both i) and ii) to stabilize the target scaffold while simultaneously providing variability in the functional hot spots.

A set of the amino acids from the functional variants should be included at the functional hot spots according to their frequencies in the variant profile because they are evolutionarily selected or optimized. Furthermore, the variants at each position can be filtered or prioritized to include other potentially beneficial mutants or exclude potentially undesirable mutants to meet the computational and experimental constraints.

Although the variant profile is informative on the preferred amino acid residues at each position and specific mutants in a preferred order, unmodified, it embodies an enormous number of recombinants. Some filtering using frequency cutoff can reduce the combinatorial sequences that need to be evaluated by computational screening or targeted directly by experimental libraries.

A structure-based scoring was applied to screen the hit library and its combinatorial sequences that form a hit variant library. Side chains of V_H CDR3 of the parental anti-VEGF antibody were substituted by rotamers of corresponding amino acid variants from the hit variant library at each residue position. The conformations of rotamers were built and optimized by using the program SCWRL® (version 2.1) using backbone-dependent rotamer library (Bower MJ, Cohen FE, Dunbrack RL (1997) JMB 267, 1268-82).

The scoring was done by searching the optimal rotamers and minimizing the energy by 100-200 steps using the Amber94 force field in CONGEN [Bruccoleri and Karplus (1987) Biopolymers 26:137-168] in the presence and absence of the structure of the antigen VEGF. The energy scores of an anti-VEGF variant library based on the calculated scores with and/without VEGF antigen indicate there are a large number of sequences for various variant libraries with higher scores than the parental sequences.

A refined custom scoring function that includes sidechain entropy, nonpolar solvation energy and electrostatic solvation energy can be used to gauge the simple functions used for scoring. Three energy terms were calculated: sidechain entropy, nonpolar solvation energy and electrostatic solvation energy and the backbone entropy for loops as well.

The nonelectrostatic solvation energy is made proportional to the molecular surface, as calculated by the GEPOL93 algorithm, with the scaling constant of 70 cal/mol/A^2 (Tunon I, Silla E, Pascual-Ahuir JL (1992) Prot Eng 5, 715-716) using GEPOL (Pascual-Ahuir JL, Silla E (1993) J Comput Chem 11, 1047-1060) command as implemented in CONGEN.

The electrostatic solvation energy is calculated using the finite-difference PB (FDPB) method as implemented in UHBD program (Davis ME, Madura JD, Luty BA, McCammon JA (1991) Comput Phys Commun 62, 187-197). The focusing method is

used for the region surrounding the mutation. An automated protocol generates three grids: coarse, fine, and focus grids. The grid units are 1.5, 0.5, and 0.25 angstroms, respectively. The focusing grid is a cubic grid that spans the Cartesian volume occupied by the mutated residues. The fine grid is a cubic grid that spans the entire volume of the protein or the complex. The coarse grid is a cubic grid that is set to approximately twice the size of the fine grid in each axis and covers approximately 8 times the volume of the fine grid. The coarse grid serves to account for the long-range solvent effects and sets the boundary conditions for the fine grid. Similarly, the fine grid accounts for the electrostatic contributions of the protein interior and sets the boundary condition for the focus grid. The focus grid accounts for finer details of the localized effects due to the mutation. The dielectric constants for the protein interior and exterior are set to 4 and 78, respectively. Temperature is set to 300 Kelvin and ionic strength is set to 150 mM. Maximum iteration is set to 200. The calculations are repeated with a uniform dielectric so that both the interior and exterior dielectrics are set to 4 and the difference between the two energies is computed. The latter calculations represent the energies due to bringing the charges onto the grids.

The custom scoring function or the molecular mechanics energy using Amber94 forcefield in CONGEN plus the solvation terms from PB in UHBD used here is similar to MM-PBSA or MM-GBSA. The energy function shows better agreement with experimental data (Sharp KA. (1998) Proteins 33, 39-48; Novotny J, Bruccoleri RE, Davis M, Sharp KA (1997) J Mol Biol 268, 401-411), especially when structure ensembles by molecular dynamics calculations are used to provide more accurate methods to score sequence and its variants based on the ensemble averages of the energy functions (Kollman PA, Massova I, Reyes C, Kuhn B, Huo SH, Chong LT, Lee M, Lee TS, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan P, Case DA, and Cheatham TE (2000) Acc. Chem Res. 33, 889-897).

The variant profile from the hit variant library as described above was filtered in order to reduce the potential library size while maintaining most of the preferred residues. A strategy that selects top sequences based on favorable score and/or the presence of residues likely to participate in favorable interactions was employed to identify a cluster or clusters of amino acid sequences for the nucleic acid library design. As described

above, a cluster of computationally screened sequences for V_H CDR3, CDR2 and CDR1, respectively, was chosen for further experimental test in vitro.

The hit variant library constructed above was targeted with a single degenerate nucleic acid library. The nucleic acid sequence profile was built using the optimal codons for specific organism that are combined with minimal sizes. A degenerate nucleic acid library was synthesized by incorporating a mixture of bases into each degenerate position. As a result of the combinatorial effect of the synthesis, this degenerate nucleic acid library encodes an expanded amino acid library.

Variants at each position for VH and VL are shown in the VH and VL amino acid position variant profile, respectively. Detailed results for V_H and/or V_L CDR3, CDR2, and CDR1, FR123 were summarized in anti-VEGF VH and VL Tables. These experiments demonstrated that by using the methods of the present invention, antibodies could be selected, not only with diverse sequences and phylogenic distances, but also with improved binding affinity to bind to the target VEGF antigen.

Figure 1A is an amino acid positional variant profile (AA-PVP) for V_L of anti-VEGF antibodies designed using the methodology described above. The AA-PVP of the V_L chain uses the humanized V_L (SEQ ID NO:1) by Baca et al. (1997) J Biol Chem 272: 10678-10684 as a reference sequence. The dot indicates the same amino acid as reference. CDRs are indicated in bold letters underlined.

Figure 1B is an amino acid positional variant profile (AA-PVP) for V_H of anti-VEGF antibodies designed using the methodology described above. The AA-PVP of the V_H chain uses the humanized V_H (SEQ ID NO:55) by Baca et al. (1997), *supra* as a reference sequence. The dot indicates the same amino acid as reference. CDRs are indicated in bold letters underlined.

The antibody libraries that were designed *in silico*, based on a lead sequence of the parental anti-VEGF antibody by using the methods described above were tested for their ability to bind to the antigen, VEGF, by using a novel phage display system.

2) Generation of anti-VEGF antibody libraries for framework optimization

A murine monoclonal antibody was found to block VEGF-dependent cell proliferation and slow the tumor growth in vivo (Kim KJ, Li B, Winer J, Armanini M, Gillett N, Phillips HS, Ferrara N (1993) Nature 362, 841-844). This murine antibody was humanized (Presta LG, Chen H, O'Connor SJ, Chisholm V, Meng YG, Krummen L, Winkler M, Ferrara N (1997) Cancer Res. 57, 4593-4599; Baca M, Presta LG, O'Connor SJ, Wells JA (1997) J Biol Chem 272, 10678-10684) using random mutagenesis at some key framework positions following grafting of antigen-binding loops. Typically, after rounds of site-directed mutagenesis and selection, humanized antibodies are generated by replacing a human or concensus human framework with non-human amino acids from the parental non-human antibody at certain pre-determined key positions. These humanized antibodies will usually bind to its cognate antigen of its parental antibody with the reduced affinity relative its parental antibody (about 6-fold weaker for humanized anti-VEGF relative its parental murine antibody, see Baca M, Presta LG, O'Connor SJ, Wells JA (1997) J Biol Chem 272, 10678-10684, and 2-fold weaker for another version of the humanized anti-VEGF, see Presta LG, Chen H, O'Connor SJ, Chisholm V, Meng YG, Krummen L, Winkler M, Ferrara N (1997) Cancer Res. 57, 4593-4599; Baca M, Presta LG, O'Connor SJ, Wells JA (1997) J Biol Chem 272, 10678-10684). This loss of binding affinity would be recovered by using affinity maturation in CDRs (Chen Y, Wiesmann C, Fuh G, Li B, Christinger HW, McKay P, de Vos AM, Lowman HB (1999) J. Mol Biol 293, 865-881).

Using present inventive methods described, we have discovered several humanized frameworks that are several-folder higher in binding affinity (3.3-fold for hAB2, 4-fold for hAB3 and 2-fold for hAB4) upon framework optimization than the reported humanized anti-VEGF antibody sequence (see hAB1 in **Figure 2** for the humanized anti-VEGF antibody framework reported in the literature (Presta LG, Chen H, O'Connor SJ, Chisholm V, Meng YG, Krummen L, Winkler M, Ferrara N (1997) Cancer Res. 57, 4593-4599). Because the reported humanized anti-VEGF antibody is ~ 2 times weaker than its corresponding murine antibody, our humanized antibodies (hAB2 and hAB3) should have ~2-fold higher binding affinity upon humanization than the corresponding murine antibody. Also, we can achieve such improvement by either

increasing the on-rate (K_{on}) or decreasing the off-rate (K_{off}) or both of the humanized antibodies relative to the reported humanized antibody (hAB1) (see **Figure 2**).

The amino acid sequence of the framework fr123 region of the murine anti-VEGF antibody is:

EIQLVQSGPELKQPGETVRISCKASWVKQAPGKGLKWMGRFTFSLETSASTAYL QISNLKNDDTATYFCA.

This sequence is therein after referred to as fr123 of "murine anti-VEGF antibody", *see* Presta LG, Chen H, O'Connor SJ, Chisholm V, Meng YG, Krummen L, Winkler M, Ferrara N (1997) Cancer Res. 57, 4593-4599). The relatively constant framework 4 can be designed if desired using the same approach. The framework and CDRs are designated according to the Kabat criteria (Kabat EA, Redi-Miller M, Perry HM, Gottesman KS (1987) Sequences of Proteins of Immunological Interest 4th edit, National Institutes of Health, Bethesda, MD), although other classification can be used also. Also, separate segment of framework FR1 or FR2 or FR3 and FR4 can be designed individually and pasted together if desired. The combination of CDRs and FRs can be designed simultaneously by designing each segment or combinations of segments used the approach described here. The CDRs are the same as in SEQ283 from the murine anti-VEGF or those designed using the approach described here. However, different CDRs can be also designed and used in combination with the designed FR123 libraries. The variant profiles for the hit library are generated using the human V_H germline sequences based on the lead sequence of V_H FR123 of the murine anti-VEGF antibody and are filtered by using certain cutoff values. The variant at each position can be ranked based on its structural compatibility with the antibody structure using total energy or other scoring terms. Some reference amino acids are found to be favorable at certain positions based on their total energy or specific packing, although their occurrence frequency is very low. The variant profiles for the hit library can be generated and refined using the Kabat-derived human VH sequences based on the lead sequence of V_H FR123 of the murine anti-VEGF antibody. The filtered variant profile can be further screened computationally to reflect the ranking order of the structural compatibility if only the antibody structure is used. Although the human vs non-human sequences differ

in many positions across the entire chain for V_H, the amino acid libraries used in other humanization approach are focused on randomization at a few positions flanking grafted CDRs, whereas in a preferred embodiment, the humanization library targets various positions across both VH and VL chains with a few mutants at those positions for the starting anti-VEGF antibody.

In a preferred embodiment, each motif such as frameworks FR1, FR2, FR3 and FR4 or its combination such as FR123 of the antibody can be targeted using a modular in silico evolutionary design approach. It has been understood that there are only a limited number of conformations (called canonical structures) for each motif or its combination. These structural features of an antibody provide an excellent system for testing the evolutionary sequence design by using structured motifs at various regions of an antibody based on the extensive analysis of antibody structures. These structure and sequence conservation are observed across different species. In fact, the scaffolding of antibodies, or the immunoglobin fold, is one of the most abundant structure observed in nature and is highly conserved among various antibodies and related molecules.

In a preferred embodiment, the method can be also used to design antibody framework using sequence-based approach or structure ensembles that contain the induced structure changes in CDRs. Using murine anti-VEGF antibody framework as the lead protein and its V_H FR123 as the lead sequence, digital libraries of V_H FR123 were constructed.

In a preferred embodiment, a hit library was constructed by searching and selecting hit amino acid sequences using V_H FR123 as the lead sequence. Variant profile was built to list all variants at each position based on the hit library and filtered with certain cutoff value to reduce of the size of the resulting hit variant library within computational or experimental limit. Variant profiles were also built in order to facilitate i) the sampling of the sequence space that covers the preferred region in the fitness landscape; ii) the partitioning and synthesis of degenerate nucleic acid libraries that target the preferred peptide ensemble sequences; iii) the experimental screening of the antibody libraries for the desired function; and iv) the analysis of experimental results with feedback for further design and optimization.

The lead structural templates were obtained from the available X-ray structures of the complexes formed between VEGF and anti-VEGF antibodies. The complex structure of VEGF and parental anti-VEGF antibody is designated as 1BJ1, and that formed between VEGF and matured anti-VEGF antibody 1CZ8. The results from 1CZ8 structural template were similar to those from 1BJ1 in the relative ranking order of the scanned sequences. The modeled structure or structure ensemble or ensemble average can be also used for screening sequences. The lead sequence for V_H FR123 (SEQ283) is taken from the murine anti-VEGF antibody according to Kabat classification. The HMM built using the single lead sequence was calibrated and used to search human heavy chain germline sequence database and/or human sequence database (including human germlines and humanized sequences) derived from Kabat database (Johnson, G and Wu, TT (2001) Nucleic Acids Research, 29, 205-206). All sequence hits that are above expectation value or E-value are listed and aligned using HAMMER 2.1.1 package. After removing the redundant sequences from the hit list, the remaining hit sequences for the lead HMM form the hit library. The sequence identities of the hit sequences from the human V_H germline ranges from 40 to 68% of the lead sequence, whereas the corresponding sequence identities of the hit sequences from human immunoglobin sequences derived from Kabat database (the database are parsed to fr123 fragment in order to increase the sensitivity of the search and their relative ranking) (other database would be used if the contain the immunoglobin sequences of human origins) ranging from ~30 to 75%. The evolutionary distances between the hits can be analysed by using the program TreeView1.6.5 (http://taxonomy.zoology.gla.ac.uk/rod/rod.html).

The AA-PVP tables give the number of occurrence of each amino acid residue at each position. There are some differences between the hit sequences from the human VH germlines and those from Kabat-derived human V_H sequences: amino acids of non-human origins resulting from amino acids that are structurally important to stabilize the scaffold of the target antibodies etc. This filtered variant profile can be further screened computationally to reflect the ranking order of the structural compatibility if only the antibody structure is used. In short, using different database of human origin for framework optimization would provide diverse but powerful choices of amino acids for framework optimization including humanization with improved binding affinity and

stability. With the increase in our knowledge in developing therapeutic antibodies, more and more antibody sequence data will be accumulated and guide our design using present invention. No prior assumption is needed to assume the key positions and amino acids associated with those positions. Because this information is revealed automatically using present inventive method, it will become better defined with increase in their occurrence in database as more data are accumulated. Variants can be re-profiled or prioritized to include other potentially beneficial mutants using structure-based criteria. The structure-based energy scoring provides another way to re-profile the occurrence of variants at each position for the hit variant library which was originally built based on profiling of sequences selected from protein databases. Some filtering using frequency cutoff can reduce the combinatorial sequences that need to be evaluated by computational screening or targeted directly by experimental libraries. Even with the cutoff applied to the variant profile, there is still a large number of combinatorial sequences that needs to be scored and evaluated in the final sequences for experimental screening.

A structure-based scoring is applied to screen the hit library and its combinatorial sequences that form a hit variant library. Side chains of V_H FR123 of the anti-VEGF antibody in 1CZ8 or 1BJ1 were substituted by rotamers of corresponding amino acid variants from the hit variant library at each residue position by using the program SCWRL® (version 2.1) using backbone-dependent rotamer library (Bower MJ, Cohen FE, Dunbrack RL (1997) JMB 267, 1268-82). The scoring was done using energetic terms or their combination with a scaling factor for each term after the structure is optimized using the optimal rotamers and energy minimization using the Amber94 force field in CONGEN [Bruccoleri and Karplus (1987) Biopolymers 26:137-168] in the presence and absence of the structure of the antigen VEGF.

The designed framework VH fr123 have good structural compatibility with the structure relative to the murine reference. The human-like features of the framework optimization as defined partly by its database were also gauged using the phylogenetic distance analysis of the designed sequences with those from human germlines or humanized frameworks.

The variant profile from the hit variant library as described above was filtered in order to reduce the potential library size while maintaining most of the preferred residues

obtained from a hit variant library after eliminating amino acids with occurrences lower than the cutoff value and/or by screening sequences based on their compatibility with the structural scaffolding.

The hit variant library constructed above was targeted with a degenerate oligonucleotides. The degenerate nucleic acid library constructed above was cloned into a phage display system and the phage-displayed antibodies (ccFv) were selected based on their binding to immobilized VEGF coated onto 96-well plates. The library was installed into a phage display vector pABMD12 in which the V_H of anti-VEGF was replaced by the library. As a result, V_L and a variety of V_H generated from the library would pair to form a functional ccFv of anti-VEGF. The phage display library was then used for further panning against immobilized VEGF protein antigen.

In order to generate a library that can cover such a wide range of scattered distribution of degenerative positions, multiple overlapping degenerative DNA oligonucleotides were synthesized with degenerative positions at the sites where the library was designed. The assembly process consisted of two PCR reactions, assembly PCR, and amplification PCR. The assembly oligos were designed with 35-40mers and overlapped by 15-20 bases with melting temperature of about 60°C by average. One additional pair of amplification oligo primers (Amp93 and Amp94) was created for final amplification of the designed products. Accordingly, the assembly PCR includes: equal amount of the assembly oligo primers in a final total concentration of 8 uM, dNTP of 0.8 uM, 1x pfu buffer (Strategene), and 2.5 units of pfu turbo (Strategene). The thermal cycle was performed as follows: 94°C x 45", 58°C x 45", 72°C x 45" for 30 cycles and a final extension of 10 minutes at 72°C. The PCR product mix was diluted 10 folds and used as the template for the amplification PCR in which all reagents were remained the same except for addition of the amplification primers at the final concentration of 1 uM. The thermal cycle was performed as follows: 94°C x 45", 60°C x 45", 72°C x 45" for 30 cycles and a final extension of 20 minutes at 72°C. The final product (the VH library) was purified, digested with HindIII and StyI, and finally subcloned into vector pABMD12 to replace the original murine V_H. The library was used to electrically transform TG1 competent cells, which were in turn amplified and rescued by helper

phage KO7 (Amersham) before production of phages of the library at 30°C overnight according to standard procedure.

To screen the library constructed described in the above example, purified homodimeric VEGF protein (Calbiochem) was diluted in designated concentration in coating buffer (0.05 M $NaHCO_3$, pH 9.6) and immobilized on Maxisorb wells (Nunc) at 4°C overnight. The coated wells were then blocked in 5% milk at 37°C for 1 hr before phage library diluted in PBS was applied in the wells for incubation at 37°C for 2 hrs. The incubation mix also routinely contained 2% milk to minimize nonspecific binding. At the end of the incubation, the wells were washed and the phages bound were subsequently eluted by 1.4% triethylamine before infecting TG1 cells followed by rescue by KO7 helper phage for amplification. To amplify the phages, infected and rescued TG1 cells were then grown at 30°C overnight in the presence of carbenicilline and kanamycin before phage library was harvested. The phages amplified were used as the input library for the next round of panning. Meanwhile, individual clones from 5th panning and on were randomly sampled for phage ELISA, in which specific binding to immobilized VEGF would be confirmed, and demonstrated 100% positives from the 5th to 7th pannings. Finally, isolated clones grown on plates of 2xYT/carbenicilline (100 ug/ml)/kanamycin (70 ug/ml) were sampled for sequencing beginning from the 5th panning (P5) to define the hit positions and hit sequences against the design.

Figure 1C shows amino acid sequences of full length V_L (as compared with that of the humanized V_L (SEQ ID NO:1) as in Baca et al. (1997) J Biol Chem 272:10678-10684, and mouse anti-VEGF monoclonal antibody (SEQ ID NO:284) as in Kim et al. (1993) Nature 362:841-844, V_L/CDR, and V_L/FR of selected anti-VEGF antibodies.

Figure 1D shows amino acid sequences of full length V_H (as compared with that of the humanized V_H (SEQ ID NO:55) as in Baca et al. (1997) J Biol Chem 272:10678-10684, that of affinity-matured V_H (SEQ ID NO:56) as in Chen et al. (1999) J. Mol. Biol. 293:865-881), and that of mouse anti-VEGF monoclonal antibody (SEQ ID NO:283) as in as in Kim et al. (1993) Nature 362:841-844, V_L/CDR), V_H/CDR, and V_H/FR of selected anti-VEGF antibodies.

For the purpose of direct comparison with anti-VEGF antibodies generated by others, the humanized V_L (SEQ ID NO:1) disclosed in the present invention is the same

as the V_L of antibody Y0317 described in Chen et al. (1999) J. Mol. Biol. 293:865-881; and V_H (SEQ ID NO:55) disclosed in the present invention is the same as the V_L of antibody Y0192 described in Chen et al. (1999), supra. The humanized and affinity matured V_H (SEQ ID NO:56) disclosed in the present invention is the same as the V_H of antibody Y0317 described in Chen et al. (1999), supra.

The selected optimized V_H frameworks also cluster together with the humanized V_H sequence, very close in phylogenetic distance to the human germline V_H3 family, while the murine V_H framework is very distant from the optimized V_H frameworks and human germlines. The phylogenic analysis of the hit sequences against the entire human immunoglubin repertoire of V_H suggests that they are indeed most closely related to human germline family III.

This supports the conclusion that the present inventive method in designing optimized frameworks with fully human or human-like sequences of the optimized antibodies, depending on the fine balance between human-like and compatibility with structure template or templates from ensemble structure or structure average.

Using our inventive methods described, we have discovered numerous heavy chain (e.g., SEQ ID NO: 70, SEQ ID NO: 67 and SEQ ID NO: 75) humanized frameworks with higher binding affinity upon framework optimization than the parental or reference anti-VEGF antibody sequence (SEQ ID NO: 55). This improvement comes mainly from a larger increase in the on-rate (K_{on}) and small decrease in the off-rate (K_{off}) by framework humanization alone. **Figure 2A** shows the affinity data of 5 antibodies, parental antibody (hAB1) and the optimized frameworks (hAB2, hAB3, hAB5) of anti-VEGF antibody selected from designer libraries using BIAcore biosensor. The measurement was performed by measuring the change of SPR units (y-axis) vs time (x-axis) when a purified antibody binds its antigen (VEGF) immobilized on the CM5 biochip at 25°C. Two humanized frameworks hAB2 and hAB3 are ~4-folder higher in binding affinity (in single chain format) upon framework optimization than the parental/reference anti-VEGF antibody sequence reported in the literature (Presta LG, Chen H, O'Connor SJ, Chisholm V, Meng YG, Krummen L, Winkler M, Ferrara N (1997) Cancer Res. 57, 4593-4599), these two humanized antibodies should have ~2-fold higher binding affinity upon humanization than the corresponding murine antibody.

2. Selection of Candidate Antibodies

1) ccFv—a heterodimeric coiled-coil stabilized antibody

The present invention provides a new strategy to stabilize V_H and V_L heterodimer. A unique heterodimerization sequence pair was designed and used to create a Fab-like, functional artificial Fv fragment ccFv (US200030027247A1). This sequence pair specifically forms a coiled-coil structure and mediates the functional heterodimerization of $GABA_B$-R1 and $GABA_B$-R2 receptors. Each of the heterodimeric sequence pair was derived from the coiled coil domains of heterodimeric receptors $GABA_B$ R1 and R2, respectively. For the purpose of engineering a heterodimer of V_H and V_L of an antibody, the pair of sequences GR1 and GR2, are fused to the carboxy-terminus of V_H and V_L fragment, respectively. Thus, the functional pairing of V_H and V_L, ccFv (coiled coil Fv), is mediated by specific heterodimerization of GR1 and GR2. Recombinant ccFv antibody fragments were expressed with a molecular weight 35 kDa.

V_H and V_L sequences of an anti-VEGF antibody was cloned into a vector that expressed two fusion proteins: V_H-GR1 and V_L-GR2-pIII fusions. The expressed V_H-GR1 and V_L-GR2-pIII fusions are secreted into periplasmic space, where they heterodimerize to form a stable ccFv antibody via the coiled-coil domain. To display ccFv antibodies on the surface of filamentous bacteriophage, the vector above was transformed into bacterial TG1 cells, which were further superinfected with KO7 helper phage.

2) Adaptor-mediated phage display system

In the conventional phage display system, a protein of interest is fused to a phage capsid protein such as pIII in order to be displayed on the surface of phage. This fusion protein will be assembled into phage particles with the wild-type phage proteins provided by a helper phage such as KO7. A new phage display system named adaptor-directed display system as described in US20030104355A1 was used for displaying various antibody fragments on the surface of phage. In general, an antibody fragment such as single chain variable fragment is carried to the surface of the phage particle by a pair of

adaptors that specifically form a heterodimer, one being fused with the displayed protein in an expression vector and the other being fused with a phage capsid protein in a helper vector. Particularly, each of the heterodimeric sequence pair, GR1 and GR2, was derived from the coiled coil domains of heterodimeric receptors $GABA_B$ R1 and R2, respectively. For the purpose of displaying an antibody fragment, the sequences GR1 was fused to the carboxy-terminus of an antibody fragment in an prokaryotic expression vector, whereas GR2 was fused to the amino terminus of the capsid protein III of bacterophage genome. The heterodimer is formed via a sequence pair specifically forms a coiled-coil structure and mediates the functional display of the antibody fragment upon rescue of the E.coli carrying the expression vector by the helper phage.

3) Preparation of candidate antibody libraries

DNA of libraries of candidate antibodies was prepared based on PCR assembly using standard PCR procedure. The DNA was then restriction digested, purified, and ligated into an appropriate vector as described above. After ligation, DNA was transformed into TG1 cells. Phages were prepared from TG1 cells by a helper phage infection. The infected TG1 cells were grown in 2xYT/Amp/Kan at 30°C overnight. The phagemid particles were precipitated by PEG/NaCl from culture supernatants, and resuspended in PBS.

4) Selection of candidate antibody libraries

The phage libraries of the candidate antibodies were used for library selection against immobilized VEGF. Purified recombinant human VEGF165 was purchased from Calbiochem (cat. No: 80054-994). The glycosylated, disulfide-linked homodimer of a 165 amino acid residue variant of human VEGF has an apparent molecular weight of 42 kD (Burke et al. Biochem. Biophys. Res. Commun. 207:348 (1995); Neufeld et al. Prog. Growth Factor Res. 5: 89 (1994); Leung et al. Science 246:1306 (1989)). In general, after incubation period of binding, unbound phages were washed away and bound phages were eluted and amplified for the next round of panning. Similar procedures used in general for phage display have been illustrated in the past (Barbas et al., Phage display: a laboratory manual, Cold Spring Harbor Laboratory Press, 2001).

For example, an aliquot of 100 ul of 2ug/ml purified human recombinant VEGF165 was first immobilized onto each well of a 96-well plate. After blocking with 5% milk in PBS, an aliquot of the library phages in 2% milk/PBS was added into the well and incubated. The phage containing solution was then discarded, and the wells were washed. Bound phages were finally eluted with 100 mM triethylamine, and were added to TG1 culture for infection. The phages prepared from infected TG1 cells were consequently used for the next round of panning. Positive clones were then confirmed by ELISA against VEGF antigen protein, in which the phages bound to the immobilized antigen were detected by incubation with HRP-conjugated anti-M13 antibody against phage coat protein pVIII. The substrate ABTS [2,2'Azino-bis(3-ethylbenzthiazoline-6-sulfonic acid)] was used for measurement of HRP activity. DNAs of those clones that were positively confirmed were sequenced. DNA sequences sampled were then translated to amino acid sequences. Selected sequences were combined to generate new variants. Selected variants were expressed to produce soluble antibody fragments for affinity evaluation.

3. Amplification of V_k and V_l from Human PBLs

The total RNA was extracted from human peripheral blood lymphocytes (PBLs), and was used as a template for first-strand cDNA synthesis by using olig-dT primer and reverse transcriptase in a standard procedure (first strain cDNA synthesis kit was purchased Roche Applied Science). The antibody light-chain variable genes were amplified from the single stranded cDNA by PCR. The PCR primers were designed and synthesized based on Kaba V-gene database (table 1). In order to amplify all 6 subfamily of Vk genes, 6 degenerate primers targeted the 5' end of the 40 germline human Vk genes and 3 degenerate primers targeted the 3' end of the 5 human Jk segment genes were used. For the PCR amplification of all 10 subfamily of Vl genes, 8 degenerate primers targeted the 5' end of the 31 germline human Vl genes and 2 degenerate primers target to the 3' end of the 4 Jl segment genes were used. The PCR amplified Vk and Vl genes were further cloned into a phagemid vector carried a humanized anti-VEGF antibody VH to generate phage display library.

Table 1: Primers for Vk and Vl amplification

Vk	ATTAATGGATCCGMCATCCRGWTGACCCAGTCTCC
	ATTAATGGATCCGATRTTGTGATGACYCAGWCTCC
	ATTAATGGATCCGAAATWGTGWTGACRCAGTCTCC
	ATTAATGGATCCGACATCGTGATGACCCAGTCTCC
	ATTAATGGATCCGAAACGACACTCACGCAGTCTCC
	ATTAATGGATCCGAAATTGTGCTGACTCAGTCTCC
Vl	ATTAATGGATCCCAGTCTGTGYTGACKCAGCC
	ATTAATGGATCCCAGTCTGCCCTGACTCAGCC
	ATTAATGGATCCTCCTATGAGCTGACWCAGCyAC
	ATTAATGGATCCTCTTCTGAGCTGACTCAGGAC
	ATTAATGGATCCCTGCCTGTGCTGACTCAGCC
	ATTAATGGATCCCAGCYTGTGCTGACTCAATC
	ATTAATGGATCCCAGSCTGTGCTGACTCAGCC
	ATTAATGGATCCAATTTTATGCTGACTCAGCCC
	ATTAATGGATCCCAGRCTGTGGTGACYCAGGAG
	ATTAATGGATCCCAGGCAGGGCTGACTCAGCC
Jk	TTAATTGCGGCCGCTTTGATYTCCASCTTGGTCCC
	TTAATTGCGGCCGCTTTGATATCCACTTTGGTCCC
	TTAATTGCGGCCGCTTTAATCTCCAGTCGTGTCCC
Jl	TTAATTGCGGCCGCTAGGACGGTSASCTTGG
	TTAATTGCGGCCGCGAGGACGGTCAGCTGGG

4. Expression of Soluble Antibody Fragments

Soluble antibody fragments in the format of single chain variable fragment (scFv) can be generated in prokaryotic (E. coli) and eukaryotic (yeast) expression systems for the purpose of biophysical analysis. Construction of scFv includes a VH fragment and a VL fragment connected by a linker of $(G_4S)_3$ as described in previous studies ((Barbas et al., Phage display: a laboratory manual, Cold Spring Harbor Laboratory Press, 2001). Expression vector used for prokaryotic expression is illustrated in Figure 7. Competent bacterial cells, e.g., BL21, were prepared and transformed with a vector that carries the antibody fragment according to methods known in the art (Sambrook, Fritsch and Maniatis, Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Press, Cold Spring Harbor, N.Y., (1989); Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, 1988). The transformed cells are cultured under conditions suitable

for protein expression. Such conditions are well known to artisans in the field and hence are not detailed herein. The expressed antibodies are harvested using conventional methods known in the art and used for further analysis. Expression in yeast was performed using Pichia expression kit purchased from Invitrogen and according to manufacturer's instruction. All antibodies were tagged with a HA-His6 tag at C-terminus, and purified by NTA and Superdex 75 columns. In order to determine the purity and expression yield of antibody fragments, 20 ul of purified proteins are analyzed by SDS-PAGE gel, and visualized by staining with Coomassie Brilliant Blue R-250.

5. Affinity Analysis of Selected Antibody Fragments

BIAcore Biosensor Assay was used to determine antibody affinity. VEGF protein (purchased from Calbiochem) was coupled to a CM-5 biosensor ship by amine coupling. After immobilization, CM-5 chips with 200 to 1000 response unites of VEGF were kept at 4°C before use. All experiments were performed at 25°C or 35°C. Each sample in PBS buffer was injected over VEGF surface at a flow rate of 20 ul/min using Biacore 1000 (Biacore AB, Sweden), and bound antibodies were removed from chip by 10 ul of glycine-HCl, PH 1.5 at the end of each cycle. Each sensorgram was recorded and normalized to a PBS base line. To determine the antibody affinity including association and dissociation rate constants, the sensorgrams were analyzed by binding curve fitting to 1:1 Langmuir binding model using BIAevalvation version 3 software.

6. Evaluation of Stability of Antibody Fragments

In order to evaluate stability of the defined antibody fragments, soluble antibodies in PBS were incubated in 4°C, 37°C, and 42°C, respectively. The antibody concentration was 0.5uM or 1uM. After indicated times of incubation, antibody solutions were set at 25°C for 1 hr, then directly injected over VEGF surface in Biacore 1000. The antibody binding RU was recorded, and used for stability analysis.